

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-251540

(43) 公開日 平成8年(1996)9月27日

(51) Int.Cl.⁶

H 0 4 N 5/93

識別記号

庁内整理番号

F I

H 0 4 N 5/93

技術表示箇所

Z

審査請求 未請求 請求項の数29 O L (全 18 頁)

(21) 出願番号 特願平7-46970

(22) 出願日 平成7年(1995)3月7日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 山田 伸

神奈川県川崎市多摩区東三田3丁目10番1

号 松下技研株式会社内

(72) 発明者 金森 克洋

神奈川県川崎市多摩区東三田3丁目10番1

号 松下技研株式会社内

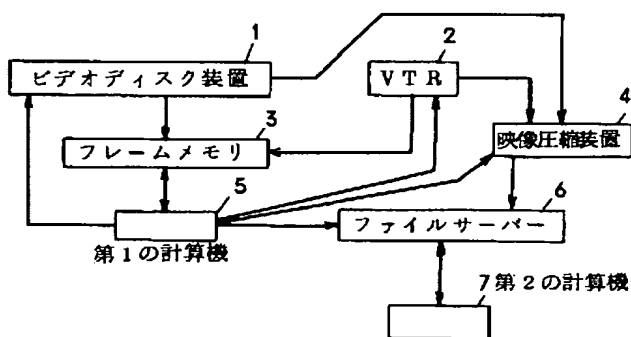
(74) 代理人 弁理士 滝本 智之 (外1名)

(54) 【発明の名称】 映像要約方法

(57) 【要約】

【目的】 映像の検索、編集、加工、早見を支援する装置に使用される、映像を要約して表示する方法に関するもので、ユーザーが要約した映像をみるときに所望の内容を把握できないという課題を解決し、被写体の動作を中心とした内容などを効率よく呈示するように、映像を自動で要約することを目的とする。

【構成】 ビデオディスク装置1又はVTR2により入力された映像信号のフレーム画像をフレームメモリ3で取り込み、第1の計算機5によりフレームメモリ3に制御信号を送ってフレーム画像を取り込み、時系列のフレーム画像を処理することで映像を要約し、そのデータをファイルサーバー6に記憶する。第2の計算機7によりユーザーの要求に応じて要約映像をファイルサーバー6から呼び出し、その映像を再生する。



【特許請求の範囲】

【請求項1】 要約対象となる映像のフレーム画像間の類似度を計算し、予め決められた時間以上にわたって類似した画像が続く時間区間である長時間類似区間を検出し、前記長時間類似区間を前記長時間類似区間以外の時間区間に比べて速く再生する映像要約方法。

【請求項2】 2枚のフレーム画像をそれぞれ複数の部分領域に分割し、前記部分領域の類似度を用いて前記2枚のフレーム画像の類似度を計算する請求項1記載の映像要約方法。

【請求項3】 要約対象となる映像のフレーム画像から、他の時間区間に比べて画面上の物体が速く動く時間区間である高速動作区間を検出し、前記高速動作区間を前記高速動作区間以外の時間区間に比べて遅く再生する映像要約方法。

【請求項4】 フレーム画像の予め決められた領域で高速動作区間を検出する請求項3記載の映像要約方法。

【請求項5】 要約対象となる映像から時系列のフレーム画像IM (Mは1以上の自然数) をサンプリングし、特定のフレーム画像IN (Nは1以上かつM未満の自然数) 中の画素PNと、次のフレーム画像IN+1の中で前記画素PNに対応する画素PN+1の周囲の画素との間の輝度差を計算し、各フレーム画像単位で前記輝度差が予め設定したしきい値 $\theta W1$ を越えるような前記画素PNを求め、前記画素PNを用いて高速動作区間を求める請求項3または請求項4記載の映像要約方法。

【請求項6】 要約対象となる映像から時系列のフレーム画像IM (Mは1以上の自然数) をサンプリングし、特定フレーム画像IN (Nは1以上かつM未満の自然数) 中の画素PNと、前記画素PNの周囲の画素との間の輝度差である第1の輝度差を計算し、前記画素PNと、次のフレーム画像IN+1中で前記画素PNに対応する画素PN+1の周囲の画素との間の輝度差である第2の輝度差を計算し、各フレーム画像単位で、前記第1の輝度差が予め設定したしきい値 $\theta W1$ 未満となり、かつ、前記第2の輝度差が前記しきい値 $\theta W1$ を越えるような前記画素PNを求め、前記画素PNを用いて高速動作区間を求める請求項3または請求項4記載の映像要約方法。

【請求項7】 動きベクトルを用いて高速動作区間を求める請求項3または請求項4記載の映像要約方法。

【請求項8】 要約対象となる映像のフレーム画像を内容ごとにまとめて複数のショットに分割し、前記ショットの内容に応じて前記ショットの表示時間長に下限を設ける請求項1ないし請求項7のいずれか記載の映像要約方法。

【請求項9】 要約対象となる映像のフレーム画像を内容ごとにまとめて複数のショットに分割し、前記ショットの境界と予め設定した周期のリズムが相関をもつ如く再生速度を決定し、リズムを想起させながら前記ショットを順次再生する映像要約方法。

【請求項10】 音を用いてリズムを想起させる請求項9記載の映像要約方法。

【請求項11】 予め設定した時間Tlenの間隔で再生を中断し、その時点が表示されているフレーム画像を予め設定した時間Tstlの間表示して (Tlen+Tstl) を周期とするリズムを発生させ、前記リズムがショットの境界と相関をもつ如く再生速度を決定する請求項9記載の映像要約方法。

【請求項12】 要約対象となる映像のフレーム画像を内容ごとにまとめて複数のショットに分割し、映像の内容に応じて再生速度決定の基準である速度上限値を求め、少なくとも一部のショットを前記速度上限値以下の速度で再生し、前記ショットの表示時間長が予め設定したしきい値である時間長上限値以下になる如く前記ショットの再生速度を決定して、ショットを順次再生する映像要約方法。

【請求項13】 要約対象となる映像のフレーム画像を内容ごとにまとめて複数のショットに分割し、前記映像の内容に応じて再生速度決定の基準である速度上限値を求め、前記ショット全体を予め設定した基準時間長Trhyで表示するために必要な再生速度が前記速度上限値以下になる場合には、前記ショットの表示時間長が前記基準時間長Trhyになる如く前記ショットの再生速度を決定し、前記ショット全体を前記基準時間長Trhyで表示するために必要な再生速度が前記速度上限値を越える場合には、前記ショットの表示時間長を前記基準時間長Trhyの2倍に設定して、前記表示時間長の前半のTrhyの間の再生速度が前記速度上限値以下になる如く前記ショットの再生速度を決定し、時系列のショットを順次再生する映像要約方法。

【請求項14】 フレーム画像間の類似度を計算し、予め決められた時間以上にわたって類似した内容が続く時間区間である長時間類似区間を検出し、前記長時間類似区間を含むショットでは前記長時間類似区間を含まないショットに比べて速度上限値を大きい値に設定する請求項12または請求項13記載の映像要約方法。

【請求項15】 2枚のフレーム画像をそれぞれ複数の部分領域に分割し、前記部分領域の類似度を用いて2枚のフレーム画像の類似度を計算する請求項14記載の映像要約方法。

【請求項16】 要約対象となる映像のフレーム画像から、他の時間区間に比べて画面上の物体が速く動く時間区間である高速動作区間を検出し、前記高速動作区間を含むショットでは前記高速動作区間を含まないショットに比べて速度上限値を小さい値に設定する請求項12または請求項13に記載の映像要約方法。

【請求項17】 フレーム画像の予め決められた領域で高速動作区間を検出する請求項16に記載の映像要約方法。

【請求項18】 要約対象となる映像から時系列のフレ

ーム画像 IM (M は 1 以上の自然数) をサンプリングし、特定のフレーム画像 IN (N は 1 以上かつ M 未満の自然数) の中の画素 PN と、次のフレーム画像 $IN+1$ の中で前記画素 PN に対応する画素 $PN+1$ の周囲の画素との間の輝度差を計算し、各フレーム画像単位で前記輝度差が予め設定したしきい値 $\theta W1$ を越えるような前記画素 PN を求め、前記画素 PN を用いて高速動作区間を求める請求項 16 または請求項 17 記載の映像要約方法。

【請求項 19】 要約対象となる映像から時系列のフレーム画像 IM (M は 1 以上の自然数) をサンプリングし、特定フレーム画像 IN (N は 1 以上かつ M 未満の自然数) の中の画素 PN と、前記画素 PN の周囲の画素との間の輝度差である第 1 の輝度差を計算し、前記画素 PN と、次のフレーム画像 $IN+1$ 中で前記画素 PN に対応する画素 $PN+1$ の周囲の画素との間の輝度差である第 2 の輝度差を計算し、各フレーム画像単位で、前記第 1 の輝度差が予め設定したしきい値 $\theta W1$ 未満となり、かつ、前記第 2 の輝度差が前記しきい値 $\theta W1$ を越えるような前記画素 PN を求め、前記画素 PN を用いて高速動作区間を求める請求項 16 または請求項 17 記載の映像要約方法。

【請求項 20】 動きベクトルを用いて高速動作区間を求める請求項 16 または請求項 17 記載の映像要約方法。

【請求項 21】 要約対象となる映像のフレーム画像を複数のショットに分割し、時系列の前記ショットの間の相関を用いて類似している前記ショットを統合してショット群とみなし、前記ショット群から複数の時系列のフレーム画像である部分動画像を選択し、前記部分動画像を順次再生する映像要約方法。

【請求項 22】 ショット群に含まれるすべてのショットから共通する特徴を求め、前記特徴に基づいて部分動画像を加工し、加工後の部分動画像を順次再生する請求項 21 記載の映像要約方法。

【請求項 23】 ショット群に含まれるすべてのショットから共通する特徴を求め、前記特徴に基づいて部分動画像を加工し、加工後の部分動画像の先頭のフレーム画像を縮小して画面上に一覧表示する請求項 22 記載の映像要約方法。

【請求項 24】 類似背景、類似被写体が撮影されている時系列のショットを検出し、ショット群とする請求項 21 ないし請求項 23 のいずれか記載の映像要約方法。

【請求項 25】 ショットの中から選択したフレーム画像である代表時空間画像を処理して、前記代表時空間画像を動きによる変化の影響を受ける領域である動領域と前記動領域以外の領域である静止領域とに分け、それぞれの領域ごとに色のヒストグラムを作成し、前記色のヒストグラムを用いてショットの間の相関を求める請求項 21 ないし請求項 24 のいずれか記載の映像要約方法。

【請求項 26】 ショット群の中で、動領域の中に現れる色である動構成色の共通部分と、静止領域の中に現れ

る色である静止構成色の共通部分とを用いて部分動画像を加工する請求項 25 に記載の映像要約方法。

【請求項 27】 時間長がほぼ等しく、かつ、時間的に連続するショットを検出し、ショット群とみなす請求項 21 ないし請求項 23 のいずれか記載の映像要約方法。

【請求項 28】 ビデオカメラで時間的に連続して撮影された部分をショットとみなす請求項 8 ないし請求項 27 のいずれか記載の映像要約方法。

【請求項 29】 シナリオの一つのシーンをショットとみなす請求項 8 ないし請求項 27 のいずれか記載の映像要約方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、映像の検索、編集、加工、早見などを支援する方法に係り、特にビデオテープやビデオディスクに格納された映像を要約して、再生または表示をする映像要約方法に関するものである。

【0002】

【従来の技術】 近年、計算機などを応用して、映像の検索、編集、加工、早見などを支援する方法に関する研究が盛んになってきている。その一例として、ビデオ、映画の映像の内容を短時間に把握するために、映像の一部分や映像全体を短時間で再生する映像要約方法がある。

【0003】 従来の方法としては、早送り再生法、特開平 4-237284 に記載の映像内容圧縮表示処理方法、ショット毎可変速再生法（大辻、外村、「動画像高速ブラウジングの主観評価」、1993 年電子情報通信学会春季大会、SD-9-3）、ショット毎ラッシュ再生法が提案されている。なお、ショットとは、映像編集などの分野でしばしば使われる映像の単位であり、映像内容の最小単位に近い。

【0004】 従来の早送り再生法は、フレーム画像を一定時間間隔で間引いて再生する映像要約方法である。

【0005】 映像内容圧縮表示方法は、時系列のフレーム画像間の変化量を用いて、各フレーム画像の表示重要度を判定し、重要度の高いフレーム画像ほど表示に多くの時間を割く映像要約方法である。実用化のためには、フレーム画像の表示重要度を判定する部分が重要である。

【0006】 隣接フレーム画像間の画素単位の輝度変化量は、画面内の物体の動きに敏感である。そこで、動きの少ないところを飛ばし、動きの多いところはゆっくりみたい場合に用いる表示重要度判定方法として、この画素単位の輝度変化量を上記表示重要度とみなす方法が提案されている。

【0007】 一方、隣接フレーム画像間のフレーム単位の輝度変化量は、ショット内での物体の移動には比較的鈍感で、ショットが変化するときのように、フレーム全体の輝度分布傾向が変化するような場合に大きな値が出る。そこで、ショットの変化を注意して見たい場合に用

いる表示重要度判定方法として、このフレーム単位の輝度変化量を上記表示重要度とみなす方法が提案されている。

【0008】従来のショット毎可変速再生法は、各ショットの表示時間を一定にするように再生速度を制御しながら、映像を再生する映像要約方法である。

【0009】また、ショット毎ラッシュ再生法は、ショットの先頭部分を標準速度で次々に再生する映像要約方法である。

【0010】最後に、ショットについて説明を補足しておく。一つのビデオカメラで時間的に連続して撮影された部分をショットと呼ぶ。ショットは、前述したように、映像内容の最小単位に近い。また、編集でつないだ部分、ビデオカメラの撮影を中断した部分が「ショットの変化」となる。

【0011】ただし、パン、ズームなどのカメラ操作によって映像内容が変化する部分を、例外として、「ショットの変化」とみなす場合もある。この場合には、例外を考慮しない場合に比べて、ショットが映像内容の最小単位に近くなる。

【0012】映像を自動的にショットに分割する方法としては、映像変化モデル法（山田、藤岡、金森、松島、坂内、「編集効果を含む映像のシーンチェンジ検出方法」、マルチメディアと映像処理シンポジウム'94）などが提案されている。

【0013】

【発明が解決しようとする課題】しかしながら上記の早送り再生法では、映像内容に関係なく一定速度で再生するので、「主観的にみて再生速度が速い部分、遅い部分」が存在し、内容把握が困難で、かつ、ユーザーが疲れやすいという課題を有していた。

【0014】また、上記の映像内容圧縮表示方法では、表示重要度の判定方法として、隣接フレーム画像間の画素単位の輝度変化量を用いる方法と、フレーム単位の輝度変化量を用いる方法しか提案されていなかった。前者を用いた場合には、画素単位の輝度変化量が主観的な動き評価に不相当であるため、「主観的にみて再生速度が速い部分、遅い部分」が存在し、内容の把握が困難で、かつ、ユーザーが疲れやすいという課題を有していた。また、後者を用いた場合には、各ショットの先頭フレーム画像だけが表示されるため、被写体の動作を中心として内容を把握したいときに使えないという課題を有していた。

【0015】上記のショット毎可変速再生法では、ショットの時間長によって再生速度が決まるので、ユーザーが一部のショットの内容を把握できず、映像に含まれる内容の時間的な流れを知りたいときに使えないという課題を有していた。

【0016】内容の時間的な流れを知りたいときには、内容の単位である各ショットの内容を把握することが重

要であり、ショットの見落としを避けるために、ショットの変化を予測できることが重要となる。しかしショット毎可変速再生法では、「編集でつないだ部分と、ビデオカメラの撮影を中断した部分でショットが変化する」場合を除くと、ショットの変化を予測できない。同一のカメラ操作で撮影された部分を上記ショットとみなして扱うときや、シナリオの一つのシーンを上記ショットとみなして扱うときなどでは、ショットの変化を予測できず、ショットの見落としが発生するという課題を有していた。

【0017】映像では、複数のショットの組み合わせが、シーンのようなショットよりも高次の内容の単位となる。上記のショット毎可変速再生法では、すべてのショットを少しずつ再生するので、類似した内容の部分を連続して再生することになり、できるだけ異なる内容の部分を効率よく見たいときに使えないという課題を有していた。

【0018】本発明は上記従来技術の課題を解決するので、ユーザーが所望の内容を把握できる映像要約方法を提供することを目的とする。

【0019】

【課題を解決するための手段】この目的を達成するために、第1に、高速で動く物体を含む画像の時間区間を検出して高速動作区間とみなす。また、一定時間以上にわたって類似した画像が続く時間区間を検出して長時間類似区間とみなす。そして、高速動作区間を、高速動作区間以外の時間区間に比べて低速で再生し、長時間類似区間を、長時間類似区間以外の時間区間に比べて高速で再生する。さらに、映像を内容ごとにまとめて複数のショットに分割してから、各ショットの表示時間長に下限を設ける。

【0020】第2に、映像を内容ごとにまとめて、ショットに分割してから、予め設定した周期のリズムを想起させながら、時系列のショットを次々に再生する。ただし、各ショットの表示時間長に上限を設ける。また、次の2つの条件の少なくとも一方を満足するように再生速度を決定する。

【0021】リズム条件：ショットの境界と、予め設定した周期のリズムが相関をもつ。

【0022】内容条件：「再生速度決定の基準である速度上限値以下の速度で再生した部分は必ず内容を把握できる」ということを保証するように速度上限値が設定されているとき、各ショットの少なくとも一部分の再生速度が速度上限値以下になる。

【0023】第3に、要約対象の映像をショットに分割する。そのあとで、時系列のショットの間の相関を用いて、類似ショットを統合してショット群を作成する。そして、各ショット群から一つずつの部分動画像を選択し、部分動画像を次々に再生する。

【0024】

【作用】これらの構成によって、第1に、高速で動く物体が存在しなくなり、かつ、各内容が一定時間以上表示されるので、「主観的に再生速度が速い部分」が消滅する。また、映像が短時間で次々に変化するので、「主観的に再生速度が遅い部分」が消滅する。従って、従来法に比べてユーザーの疲労が軽減される。また、この方法は、主観的にみて、再生速度が許容範囲内に入るようにしながら映像全体を再生するので、被写体の動作を中心とした内容を把握することができる。

【0025】第2に、同じ内容の部分をショットにまとめ、それぞれを短時間で再生する。従って、リズム条件を満足するように映像を要約するとき、ショットの変化を予測できるので、すべてのショットを見落とさずに見ることができる。

【0026】一方、内容条件を満足するように映像を要約するとき、ユーザーは、すべてのショットにおいて、少なくともその一部分の内容を把握することができる。同じ内容の部分をまとめたものがショットになっているので、すべてのショットの内容を把握できることになる。

【0027】なお、この手段は、同じ内容の部分をまとめてショットを作成し、これらを次々に再生するので、映像に含まれる内容の時間的な流れを知りたい場合に用いる。ただし、リズム条件を満足しない場合には、ショットの変化を予測できないので、ショットの見落としが発生する。内容条件を満足しない場合には、一部のショットの内容を把握できない。リズム条件と内容条件を同時に満足することが望ましい。

【0028】第3に、時系列の類似する内容の部分をまとめてショット群を作成し、これらの一部分を次々に再生するので、できるだけ異なる内容の部分を効率よく見ることができる。

【0029】

【実施例】以下、本発明の一実施例について、図面を参照しながら説明する。

【0030】図1は、本発明の一実施例における映像要約装置の全体システム図である。図1において、1、2は処理対象となる映像（以下、処理対象映像と呼ぶ）の入力装置であって、1はビデオディスク装置、2はVTRである。また、3はビデオディスク装置1やVTR2から出力される映像信号のフレーム画像を取り込むフレームメモリである。4は、ビデオディスク装置1やVTR2から出力される映像信号を圧縮する映像圧縮装置である。5は第1の計算機であり、ビデオディスク装置1やVTR2やフレームメモリ3や映像圧縮装置4を制御する。また、フレームメモリ3に制御信号を送ってフレーム画像を取り込み、時系列のフレーム画像を処理することで映像を要約する。6は映像圧縮装置4で圧縮された映像や第1の計算機5から送られるデータやフレーム画像を記憶するファイルサーバーである。7はユーザー

の要求に応じて、要約映像、処理対象映像をファイルサーバー6から呼び出し、その映像を再生する第2の計算機である。

【0031】以上のように構成された映像要約装置について、図2に示すフローチャートを用いてその全体の動作を説明する。

【0032】手順201では、図1における第1の計算機5が、ビデオディスク装置1とVTR2とフレームメモリ3を制御しながら時系列のフレーム画像を処理することで、映像を要約する。

【0033】手順202では、第1の計算機5が、ビデオディスク装置1とVTR2と映像圧縮装置4を制御しながら、要約した映像と処理対象映像とを圧縮してファイルサーバー6に記憶する。

【0034】なお、要約映像の記憶方法については、映像を直接圧縮する方法だけでなく、他の方法が各種考えられる。例えば、図3に示すように、処理対象映像のフレーム番号を用いてファイル形式で要約映像を表現できる場合には、要約映像の代わりに、フレーム番号の情報をファイルサーバー6に記憶してもよい。また、ビデオディスク装置1とVTR2とフレームメモリ3を制御しながら、後述する各ショット群の先頭フレーム画像を取り込み、要約映像の代わりにファイルサーバー6に記憶してもよい。ただし、この場合には、記憶した先頭フレーム画像を縮小して一覧表示することが、映像の要約を呈示することに相当する。

【0035】手順203では、第2の計算機7が、ユーザーの要求に応じて、要約映像、処理対象映像をファイルサーバー6から呼び出し、その映像を再生する。

【0036】以下では、図2における手順201の具体的動作である、第1の計算機5の映像要約処理について説明する。

【0037】図4は、図1における第1の計算機5の映像要約処理の一実施例のフローチャートである。

【0038】まず、手順401では、映像要約処理に用いる様々なパラメータの初期化、しきい値の設定を実行する。また、後述する手順で「フレーム画像間の類似度」や高速動作画素や静止構成色や動構成色や静止共通色や動共通色などの特徴量を計算するときに用いるために、フレームメモリ3から処理対象フレーム画像 I_n （ n は1以上の自然数）を取り込み、このフレーム画像を記憶しておく。

【0039】次に、手順402では、処理対象となる映像が終了したかどうかを判定し、映像が終了した場合には、映像要約処理を終了する。映像が終了していない場合には、手順403に進む。

【0040】手順403では、現在の処理対象フレーム画像の次のフレーム画像 I_{n+1} を新しい処理対象フレーム画像 I_n とみなす。そして、ビデオディスク装置1やVTR2に制御信号を送って、更新後の処理対象フレー

ム画像 I_n を再生し、フレームメモリ 3 からフレーム画像 I_n を取り込む。

【0041】手順 404 では、一つのビデオカメラで時間的に連続して撮影された部分であるショットの末尾を検出する。ショットの末尾を検出した場合には、このショットを新しい処理対象ショット SH_k (k は 1 以上の整数) とみなしてから、手順 405 に進む。ショットの末尾を検出できなかった場合には、何もせずに手順 405 に進む。なお、手順 405 に進んでから、この手順 404 が再び実行されるのは、次の処理対象フレーム画像 I_{n+1} を取り込んだあとである。従って、ショットの末尾を検出できない場合には、次処理対象フレーム画像 I_{n+1} を取り込んだ後に、ショットの末尾の検出を再度試みることになる。

【0042】なお、シナリオの一つのシーンなどをショットとみなしてもよい。したがって、手順 404 以降の説明で用いている「ショット」は、すべて「シナリオの一つのシーン」におきかえてもよい。

【0043】また、ショットの末尾の検出方法として、従来の技術で紹介した映像変化モデル法が提案されている。さらに、本実施例では、映像を自動的にショットに分割しているが、ユーザーが映像などを見ながら分割してもよい。映像が予めショットに分割されている場合には、手順 404 を省略してもよい。

【0044】手順 405 以降、手順 408 まだが、早送り映像の主観評価結果をもとにたてた次の 3 つの仮定を用いる映像要約方法であり、「主観的にみて再生速度が速い部分、遅い部分」を消滅させるように再生速度を決定する。

【0045】・長時間類似した画像が続く部分（以下、長時間類似区間と呼ぶ）で、再生速度が遅いと感じる。

【0046】・高速で移動する物体を含む画像の部分（以下、高速動作区間と呼ぶ）で、再生速度が速いと感じる。

【0047】・映像を内容ごとにまとめてショットに分割したとき、表示時間が短いショット（短時間長ショットと呼ぶ）で、再生速度が速いと感じる。

【0048】なお、この方法によって作成された要約映像を区間変速要約映像と呼ぶことにする。また、本実施例では、長時間類似区間、高速動作区間以外の時間区間を標準区間と呼ぶことにする。さらに、標準区間の再生速度を 4 倍速に設定する。

【0049】手順 405 は、長時間類似区間を検出する処理である。4 倍速早送り再生の映像が、4 フレーム間隔の画像から構成されるので、4 フレーム間隔の画像をサンプリングして、その画像間の類似度を調べること、この検出処理を実行する。手順 405 における具体的な動作を図 5 を用いて述べる。

【0050】手順 501 では、まず、処理対象フレーム画像 I_n のフレーム番号を調べる。次に、このフレーム画像が、4 フレーム間隔の画像列に含まれるかどうか判定する。4 フレーム間隔の画像だけを用いて長時間類似区間を検出するので、「画像列に含まれる」と判定された場合には、長時間類似区間を検出するために手順 502 に進む。「画像列に含まれない」と判定された場合には、手順 405 を終了して、図 4 の手順 406 に進む。

【0051】手順 502 では、処理対象フレーム画像 I_n と、上記画像列で一つ前の画像列のフレーム画像 I_{n-4} との間の類似度 $S(n, n-4)$ を計算する。ただし、要約映像を見るユーザーが主観的に無視できる動きとの相関がなくなるように、フレーム画像間の類似度の計算方法を決めた。

【0052】ここで、フレーム画像間の類似度の計算方法について簡単に述べる。 χ^2 検定法（長坂、田中、「ビデオ作品の場面変わりの自動検出法」、情報処理学会第 40 回全国大会、1Q-5、1990 年）など各種の方法が考えられるが、本実施例では、共通色比率法（山田、藤岡、金森、松島、「部分領域ごとの共通色に注目したシーンチェンジ検出方法の検討」、テレビジョン学会技術報告、1993 年 9 月、Vol.17, No.55, pp1-6）と同様の方法を用いる。

【0053】静止画像、主観的に無視できる動きだけを含む映像では、画面上の物体が 4 フレーム時間に画面上で移動する距離は、画面の幅に比べて非常に小さくなるはずである。筆者らの経験によれば、この距離は画面の幅の 2% 以下になる。そこで、図 6 に示すようにフレーム画像 I_n を部分領域 $R(j, n)$ 、（ただし、 j は 1 以上 16 以下の整数）に分割する。このとき、静止画像、主観的に無視できる動きだけを含む画像 I_n 、 I_{n-4} の対応する部分領域 $R(j, n)$ 、 $R(j, n-4)$ の間では、色のヒストグラムがほとんど変化しない。なお、本実施例では部分領域の数を 16 としたが、必ずしも 16 である必要はない。

【0054】一方、フレーム画像 I_n 、 I_{n-4} の間でショットの変化などが発生すると、 I_n 、 I_{n-4} の対応する部分領域の間では、被写体の変化して部分領域を構成する色が変化する。そこで、対応する部分領域の間で新しく出現した色、消滅した色の面積に応じて、対応する部分領域の類似度 $Sp(j, I_n, I_{n-4})$ が減少するように、この類似度を計算する。

【0055】そして、フレーム画像 I_n 、 I_{n-4} の対応する部分領域の類似度の平均を計算し、フレーム画像の類似度 $S(n, n-4)$ とみなす。すなわち、

【0056】

【数 1】

$$S(n, n-4) = \frac{1}{16} \sum_{j=1}^{16} Sp(j, I_n, I_{n-4}) \cdots \cdots (1)$$

を計算する。

【0057】手順503は、長時間類似区間の先頭フレーム番号を更新する処理である。長時間類似区間中の画

$$S(n-4, n) \geq \theta_{\text{still}}$$

)

を用いて調べる。(2)式が成立する場合に、(n-4)を長時間類似区間の途中とみなし、先頭フレーム番号を表す類似区間端Nbが設定されていなければ、Nbに、(n-4)を代入する。

【0058】手順504は、長時間類似区間の候補を検

$$S(Nb+(i-1) \times 4, Nb+i \times 4) \geq \theta_{\text{still}}, 1 \leq i \leq (n-Nb)/4 \quad \dots (3)$$

)

を用いて調べる。(3)式が成立するとき、類似区間端のフレーム画像I_{Nb}と上記処理対象フレーム画像I_nとの間を長時間類似区間の候補とみなして、手順505に進む。(3)式が成立しないとき、長時間類似区間の候補を決定できないので、図4の手順406に進む。

【0059】手順504で求めた候補には、「類似した

$$S(Nb, Nb+i \times 4) \geq \theta_{\text{ratio}}, 1 \leq i \leq (n-Nb)/4 \quad \dots (4)$$

)

を用いて調べる。次に、(4)式が成立する場合には、この候補の時間長が、長時間類似区間の最低時間長T_{still}

$$n-Nb \geq T_{\text{still}} \quad \dots (5)$$

5)

を用いて調べる。ただし、(5)式中の最低時間長T_{still}は、図4の手順401の実行時に設定しておく。

【0060】(4)、(5)式が同時に成立する場合には、手順504で求めた候補を、長時間類似区間とみなしてから、以上に述べてきた手順405を終了する。

【0061】(4)式が成立し、かつ、(5)式が成立しない場合には、処理対象フレーム画像I_n以降の画像を処理しなければ、手順504で求めた候補が長時間類似区間の一部かどうか判定できない。そこで、何もせずに手順405を終了する。

【0062】(4)式が成立しない場合には、この候補を、「ゆるやかな映像変化をする部分」とみなす。また、(2)式が成立するので、フレーム番号(n-4)以降を長時間類似区間の候補とみなす。長時間類似区間の候補の先頭フレーム番号を表す類似区間端Nbに、(n-4)を代入する。

【0063】なお、本実施例では、要約映像を見るユーザーが主観的に無視できる動きとの相関がなくなるように、フレーム画像間の類似度の計算方法を決めた。しかし、ユーザーの主観は、ユーザーによってばらつくので、ユーザーが注目する物体の動きを重視して、類似度の計算方法を決めてもよい。

【0064】また、本実施例では、(3)、(4)式を用いて「類似する画像が続く部分」を求めた。すなわ

$$|f(p, n-1) - f(p+d, n)| \geq \theta_{w1}, |d \cdot d| \leq L_{\min} \times L_{\min} \quad \dots (6)$$

)

像は互いに類似するので、時系列の画像の間の類似度S(n, n-4)がしきい値θ_{still}以上になるかどうかを、条件式

$$\dots (2)$$

出する処理である。類似区間端Nbのフレーム画像と、処理対象フレーム画像I_nとの間で、4フレーム間隔の画像の類似度が上記しきい値θ_{still}以上になるかどうかを、条件式

画像が続く部分」の他に、図7に示すような「ゆるやかな映像変化をする部分」が含まれる。手順505では、まず、長時間類似区間候補の先頭フレーム画像I_{Nb}と、他のフレーム画像との間の類似度がしきい値θ_{ratio}以上になるかどうかを、条件式

illに比べて長いかどうかを、条件式

$$\dots (5)$$

ち、時系列の画像の間の類似度と、類似区間の先頭フレーム画像を基準とした類似度とを用いた。しかし、類似区間の最後のフレーム画像を基準にする方法などを用いても同様の効果がえられるので、(3)、(4)式以外の方法を用いて「類似する画像が続く部分」を計算してもよい。また、(4)式だけを用いて「類似する画像が続く部分」を計算してもよい。

【0065】図4の手順405終了後、手順406に進む。手順406は、「高速で移動する物体を含む画像の部分」を検出する処理である。手順406における具体的な動作を図8を用いて述べる。

【0066】手順801は、処理対象フレーム画像I_nと、その直前の画像I_{n-1}を用いて、高速で動く物体上の画素を検出する処理である。この処理の原理を、図9を用いて説明する。

【0067】図9に示すように、フレーム画像I_{n-1}、I_nの間で、位置p(ただし、pはベクトル)上の物体が高速移動したとき、距離L_{min}以内ではもとの物体が既に存在しないと仮定する。このとき、物体の移動によって、I_{n-1}上で、位置pの画素の輝度f(p, n-1)と、I_n上で同じ位置pを中心とする半径L_{min}の円内の画素の輝度f(p+d, n)(ただし、dはL_{min}以下の長さをもつベクトル)との間の輝度差が、すべて予め設定したしきい値θ_{w1}以上になり、

が成立する。ただし、(6)式において、 $|a|$ は、スカラー量 a の絶対値を表し、 $(d \cdot d)$ は、ベクトル d の内積を表す。

【0068】さらに、 R_{min} を半径とする円よりも大きい物体のみを検出することで、ノイズを除去する。同一輝

$$|f(p, n-1) - f(p+D, n-1)| < \theta w1, |D \cdot D| \leq R_{min} \times R_{min} \dots (7)$$

が成立するとき、位置 p で(6)式が成立するかどうか調べる。

【0069】以上のことをふまえて、図8の手順801では、(6)、(7)式を同時に満たす画素 p を検出してから、この画素を、「高速で動く物体上の画素」とみなし、高速動作画素として登録する。

【0070】なお、処理を簡単にするために、画素間引きをする場合や、8画素×8ライン分の画素の平均を一つの画素とみなす場合などがある。このように、複数の画素の内容を一つの画素に代表させる場合には、モザイクのように、1画素あたりの面積が大きくなる。この場合には、図10に示すように、(6)式だけを満たす画素を高速動作画素とみなしてもよい。

【0071】また、高速で動く物体を検出する方法として、例えば、動きベクトルを用いた方法が考えられる。次の手順のように、動きベクトルを用いて高速動作画素を検出してもよい。

【0072】[動きベクトルを用いて、高速で動く物体を検出する手順の例]

手順1) 画像を複数のブロックに分割してから、時系列の2枚の画像を用いて、ブロックごとの動きベクトルを求める。

【0073】手順2) 動きベクトルの大きさが一定値以上になるブロックを求める。

手順3) 手順2で求めたブロックに含まれる画素を、高速動作画素とみなす。

【0074】さらに、(6)、(7)式では、画素 p と比較する画素を、特定の円に含まれる画素としたが、図11に示すように、画素 p を含む四角形に含まれる画素としてもよい。

【0075】また、ユーザーの注目する領域を予め想定しておき、想定した領域に含まれる画素の中から、高速移動画素を検出してもよい。例えば、図12に示すように、ユーザーが常時画面の中央付近に注目すると想定して、高速移動画素を検出してもよい。

【0076】手順802では、フレーム画像 I_{n-1} に含まれる高速動作画素の総数を調べる。そして、高速動作画素の総数が、ノイズを除去するために設定したしきい値 θ_m に比べて大きい値になるとき、このフレーム画像 I_{n-1} と次のフレーム画像 I_n の間を、高速動作区間候補とみなす。

【0077】早送り映像の主観評価の結果では、高速で動く物体が存在しても、一瞬しか動かなければ、主観的

度領域を同一物体とみなし、図9に示したように、 I_{n-1} 上で位置 p の画素の輝度 $f(p+D, n-1)$ (ただし、 D は R_{min} 以下の長さをもつベクトル)との間の輝度差がしきい値 $\theta w1$ 未満になる条件式

に無視される。そこで、手順804は、高速動作区間候補の中から、一瞬の動きによって候補になった区間を除外する。この処理の実現方法は、いくつか考えられるが、ここでは、その一方法について述べる。

【0078】本実施例では、4倍速早送り再生の映像が基準になっているので、映像を4フレームごとにまとめ、区間単位と呼ぶことにする。手順803では、手順802で高速動作区間候補が検出されたかどうかを判定する。「検出された」と判定した場合には、この高速動作区間候補を含む区間単位を UN (N は1以上の整数)と表記することとし、手順804に進む。「検出されない」と判定した場合には、高速動作区間候補が存在しないので、手順406を終了する。

【0079】手順804は、高速動作区間候補を含む区間単位が予め設定した下限値 N_{move} 以上連続するかどうかを判定する処理である。また、区間単位 UN と、その $(N_{move}-1)$ 個前の区間単位 $UN-(N_{move}-1)*4$ との間にあるすべての区間単位が高速動作区間候補を含むかどうか調べる。次に、「高速動作区間候補を含む」と判定されたとき、フレーム画像 $I_{N-N_{move}*4}$ と I_N の間を高速動作区間とみなす。

【0080】手順804の終了後、以上に述べてきた手順406を終了する。そして、図4の手順406の終了後、手順407に進む。

【0081】手順407は、区間変速要約映像のための映像再生速度を決定する処理である。最初に、高速動作区間の再生速度を2倍速に決定する。次に、長時間類似区間の中から、途中に高速動作区間を含まないものを取り出し、その再生速度を8倍速に決定する。さらに、再生速度が決定していない残りの区間を標準区間とみなし、その再生速度を4倍速に決定する。

【0082】なお、手順407では、高速動作区間、標準区間、長時間類似区間の再生速度を、それぞれ2倍速、4倍速、8倍速に設定したが、高速動作区間の再生速度が標準区間に比べて遅く、長時間類似区間の再生速度が標準区間に比べて速ければ、他の値に設定してもよい。また、長時間類似区間全体を1秒で再生するケースのように、表示時間を用いて再生速度を決定してもよい。

【0083】手順407で決定した再生速度にしたがって、映像を再生したとき、一部のショットで、表示時間が予め設定したショット長下限値 N_{shot} 未満になる。そこで、手順408で、再生速度を修正する。

【0084】手順408では、まず、手順404で求めた各ショットSHkに対して、再生に必要な時間NShk (k

$$NShk < Nshot \quad \dots \dots \dots (8)$$

)

を用いて調べる。(8)式が成立する場合には、ショットSHkを短時間長ショットとみなし、ショットの表示時間がショット長下限値Nshotになるように、再生速度を決定する。

【0085】なお、本実施例では、ショット長下限値を固定値としたが、テクスチャの細かさなどの内容に応じてショット長下限値を変化させてもよい。

【0086】手順408の終了後、手順409に進む。手順409は、ショットの境界と予め設定した周期のリズムが相関をもつように再生速度を決定してから、この「リズム」を想起させながら時系列のショットを次々に再生する映像要約方法を実現するための手順である。最初に、要約映像がショットの境界で予め設定した周期のリズムをもつように、再生速度を決定する。次に、リズムを保持しながら、ユーザーが各ショットの内容を把握できるように、再生速度を修正する。なお、手順409で作成された要約映像をリズム呈示要約映像と呼ぶことにする。

【0087】手順409における具体的な動作を図13

$$V(k) = LSHk / Tlen \quad \dots \dots \dots (9)$$

)

によって再生速度V(k)を決める。

【0091】手順1303では、ユーザーが処理対象ショットSHkの内容を把握できるように、処理対象ショットSHkの再生速度決定の基準である速度上限値Vmax

(k)を決定する。この速度上限値Vmax(k)の決定方法は各種考えられる。例えば、手順407で決定した再生速度を、処理対象ショットSHkの中で平均し、この値を速度上限値Vmax(k)に代入する方法が考えられる。また、予め設定した固定値をVmax(k)とみなす方法も考えられる。さらに、手順407で決定した長時間類似区間、標準区間、高速動作区間を利用した次の方法も考えられる。

【0092】[長時間類似区間、標準区間、高速動作区間を利用した速度上限値決定方法]

手順1) 長時間類似区間を含み、かつ、高速動作区間を

$$V(k) \geq Vmax(k) \quad \dots \dots \dots (10)$$

)

を用いて調べる。(10)式が成立する場合には、ユーザーに内容を把握させるために、処理対象ショットSHkの早送り部分の表示時間を延長し、再生速度を修正する。この修正処理は、手順1305で実行する。

【0096】(10)式が成立しない場合には、図15のショットSH1の部分のように、ショットSHkの早送り部分の再生速度をV(k)に決定してから、以上に述べてきた手順1309を終了する。

は1以上の整数)を求める。次に、ショット長下限値Nshot未満になるかどうかを、条件式

を用いて述べる。ショットの境界に予め設定した周期のリズムをもたせるためには、例えば、各ショットの表示時間を一定にするように、再生速度を決定すればよい。また、リズムを想起させる方法としては、ショットの境界で音をならす方法、ショットの境界で静止画像を表示する方法など、各種の方法が考えられる。

【0088】本実施例では、図14に示す例のように、ショットの先頭フレーム画像を静止時間Tstlの間表示してから、ショット全体を早送り時間Tlenで早送り表示する。このとき、(Tstl+Tlen)間隔のリズムが発生する。例えば、Tstlを0.2秒とし、Tlenを0.8秒に設定すると、1秒間隔のリズムが発生する。

【0089】図13の手順1301では、処理対象ショットSHkのフレーム数LSHk(kは1以上の整数)を求める。

【0090】手順1302では、処理対象ショットSHkの早送り部分の表示時間が、予め設定した早送り時間Tlenになるように、

含まないショットでは、標準区間だけのショットに比べて、速度上限値Vmax(k)を大きい値に設定する。例えば、8倍速に設定する。

【0093】手順2) 高速動作区間を含み、かつ、長時間類似区間を含まないショットでは、標準区間だけのショットに比べて、速度上限値Vmax(k)を小さい値に設定する。例えば、2倍速に設定する。

【0094】手順3) 高速動作区間と長時間類似区間を両方含むショットと、標準区間だけのショットでは、速度上限値Vmax(k)を同じ値に設定する。

例えば、4倍速に設定する。

【0095】手順1304では、(9)式で決定した再生速度V(k)が、手順1303で決定した速度上限値Vmax(k)以上になるかどうかを、条件式

【0097】すでに述べたように、手順1305は、ショットの内容をユーザーに把握させるための処理であり、ショットの表示時間を延長し、再生速度を修正する。

【0098】ショット境界にリズムをもたせるために、ショットの表示時間を「他のショットの整数倍」にする。ただし、各ショットを短時間で次々に再生するために、各ショットの早送り部分の表示時間を可能な限り短

くする。そこで手順1305では、(10)式を満足するショットの早送り部分の表示時間を、他のショットの2倍である $2 \times T_{len}$ に設定する。

【0099】次に、再生速度の決定方法について述べる。(10)式を満たすショットの前半部分の再生速度が、手順1303で決定した速度上限値 $V_{max}(k)$ 以下になるように修正する。すなわち、図15のショットSH2のように、前半部分の再生速度 $V_P(k)$ を速度上限値 $V_{max}(k)$ 以下の値に設定した上で、早送り部分全体の

$$V_P(K) \begin{cases} V_{max}(K), \frac{V(K)}{2} > V_{max}(K) \text{ のとき} \\ \frac{V(K)}{2}, \frac{V(K)}{2} \leq V_{max}(K) \text{ のとき} \end{cases} \dots\dots\dots (11)$$

によって決定できる。また、早送り部分後半の再生速度 $V_R(k)$ は、

$$V_R(K) = \frac{L_{SHk} - V_P(K) \times T_{len}}{T_{len}} \dots\dots\dots (12)$$

によって決定できる。(11)、(12)式の計算後、以上に述べてきた手順409を終了する。

【0103】なお、内容を把握できないショットがいくつか発生してもよい場合には、手順1302の終了後に手順409を終了してもよい。また、ショットの変化を予測できなくてもよい場合には、(9)式を用いずに再生速度 $V(k)$ を決定してもよい。

【0104】図4の手順409の終了後、手順410に進む。手順410から手順412までが、時系列のショットの間の相関を用いて、類似ショットを統合してショット群とみなし、一つのショット群から複数の時系列のフレーム画像(例えば1秒分)からなる部分動画像を選択してから、時系列の部分動画像を次々に再生する映像要約方法を実現するための手順である。最初に、時系列のショット間の相関を用いて、類似ショットを統合してショット群を作成する。次に、各ショット群から一つずつの部分動画像を選択し、ショット群の特徴を明示するように加工する。例えば、類似物体を含むショットを統合してショット群を作成する場合には、この類似物体がショット群の特徴となる。また、時間長がほぼ等しいショットを統合する場合には、この時間長がショット群の特徴となる。

【0105】要約映像を再生するときには、加工した部分動画像を次々に再生する。なお、加工を実行せずに、選択した部分動画像を直接再生してもよい。

【0106】手順410から手順412によって作成した要約映像を部分選択要約映像と呼ぶことにする。

【0107】手順410は、ショット群を作成する処理である。手順410における具体的な動作を図16を用いて述べる。

【0108】ショットは内容の単位なので、1つのショット内のフレーム画像は1つの共通の特徴をもつ。例えば、人物を追尾するショットでは、すべてのフレーム画

表示時間を $2 \times T_{len}$ にするように後半部分の再生速度 $V_R(k)$ を決定する。ただし、後半部分の再生速度 $V_R(k)$ は $V_{max}(k)$ を越えてもよい。また、リズムを保持するために、早送り部分の前半部分と後半部分の間に静止部分を入れる。

【0100】早送り部分前半の再生速度 $V_P(k)$ は、

【0101】
【数2】

【0102】
【数3】

像に人物が登場する。したがって、ショット内の一部のフレーム画像を用いて、ショットの内容を代表させることができる。本実施例では、処理を簡単にするために、内容を代表させるフレーム画像の枚数 N_{rep} を予め設定しておき、ショット内の N_{rep} 枚のフレーム画像(以下、代表時空間画像と呼ぶ)を用いて、ショットの内容を代表させる。

【0109】手順1601では、ショット SH_k を代表する画像を決定して、代表時空間画像 $I_{k,j}$ 、(k は1以上の整数、 j は1以上 N_{rep} 以下の整数)とみなす。ただし、代表時空間画像の決定方法としては、先頭部分の N_{rep} 枚を選択する方法、一定時間間隔の画像を選択する方法など、各種考えられる。また、処理量に制限がなければ、ショット内の全フレーム画像を代表時空間画像とみなしてもよい。

【0110】手順1601の終了後、代表時空間画像を用いて時系列の類似内容のショットを統合し、ショット群とみなす。類似内容のショットの例としては、類似背景のショットや類似被写体のショットなどがある。本実施例では、画面上の物体の色と動きに注目してショットを統合する。

【0111】筆者らの分析によれば、類似内容のショットの代表時空間画像では、各代表時空間画像に共通する色が存在し、かつ、共通する色をもつ物体が共通の動きをする。画面に現れる2つの物体が共通の動きをしているかどうかを調べることは難しいが、物体の動きによる変化は後述する画素変化領域、動きベクトルなどを用いて簡単に検出できる。そこで、代表時空間画像を処理して、画面上の物体の動きによる変化の影響を受ける領域(以下、動領域と呼ぶ)と、残りの領域(以下、静止領域と呼ぶ)とに分割する。このとき、類似内容のショットは、次の共通色比率条件を満足する。

【0112】共通色比率条件：代表時空間画像の静止領

域を構成する色、動領域を構成する色をそれぞれ静止構成色、動構成色とよぶとき、代表時空間画像中に共通する静止構成色、動構成色が存在し、これらの色をもつ画素が、各代表時空間画像でしきい値 θ shot 以上の割合を占める。

【0113】なお、共通色比率条件を満足する類似内容のショットの例を次に示しておく。

例1：同じ場所で撮影したショットが続く場合のような、類似背景のショット。図17に示したプールサイドの風景の2つのショットSHk-1、SHkの代表時空間画像Ik-1、1、Ik、1では、

Ik-1、1中の静止構成色Aの割合＝60%

Ik-1、1中の静止構成色Bの割合＝20%

Ik-1、1中の静止構成色Cの割合＝15%

Ik-1、1中の静止構成色Dの割合＝5%

Ik、1中の静止構成色Aの割合＝80%

Ik、1中の静止構成色Cの割合＝15%

Ik、1中の静止構成色Eの割合＝5%

となり、

共通する静止構成色A、CのIk-1、1中の割合の和＝75%

共通する静止構成色A、CのIk、1中の割合の和＝95%

となるので、共通する静止構成色が画面全体の75%以上を占める。

【0114】例2：同じ被写体のショットが続く場合のような、類似被写体のショット。図18に示した自動車の追尾の2つのショットSHk-1、SHkの代表時空間画像Ik-1、1、Ik、1では、

Ik-1、1中の静止構成色Vの割合＝40%

Ik-1、1中の静止構成色Wの割合＝40%

Ik-1、1中の静止構成色Xの割合＝10%

Ik-1、1中の静止構成色Yの割合＝10%

$$\text{HSV}k = \{c \mid \text{HS}(c, \text{SH}k) > \theta H\} \quad \dots\dots (13)$$

$$\text{HMV}k = \{c \mid \text{HM}(c, \text{SH}k) > \theta H\} \quad \dots\dots (14)$$

以下では、ショット群の先頭のショットをSHtop (topは1以上の整数) と記述する。手順1604から手順1606までは、このショットSHtop以降の時系列のショットを調べて、どこまでのショットが共通色比率条件を満たすか調べる処理である。

【0119】手順1604では、ショットSHtop～SHtop

$$\text{HSC}_{\text{top}, m} = \{c \mid c \in \text{HSV}_{\text{top}} \cap \text{HSV}_{\text{top}+1} \cap \dots \cap \text{HSV}_{\text{top}+m}\} \quad \dots\dots (15)$$

$$\text{HMC}_{\text{top}, m} = \{c \mid c \in \text{HMT}_{\text{top}} \cap \text{HMT}_{\text{top}+1} \cap \dots \cap \text{HMT}_{\text{top}+m}\} \quad \dots\dots (16)$$

手順1605では、「静止共通色HSCtop, mと動共通色HM

Ik、1中の静止構成色Zの割合＝40%

Ik、1中の静止構成色Wの割合＝35%

Ik、1中の静止構成色Xの割合＝10%

Ik、1中の静止構成色Yの割合＝15%

となり、

共通する動構成色W、X、YのIk-1、1中の割合の和＝60%

共通する動構成色W、X、YのIk、1中の割合の和＝60%

となるので、共通する動構成色が画面全体の60%以上を占める。

【0115】図16の手順1602では、代表時空間画像から動領域を検出して、代表時空間画像を動領域と静止領域に分割する。動きベクトルを用いると、動領域は図19に示すようになる。また、2枚のフレーム画像の間で物体が移動しないとき、同じ位置にある画素の輝度がほぼ等しいので、次の画素変化領域を動領域とみなすと、動領域は図20に示すようになる。

【0116】画素変化領域：代表時空間画像中のフレーム画像Ik, jと、その1フレーム時間後のフレーム画像との間で、「同じ位置にある画素の輝度差の絶対値がしきい値 $\theta W1$ 以上となる画素」の集合。

【0117】手順1603は、静止構成色、動構成色を求める処理である。以下、画素の赤色成分、緑色成分、青色成分をそれぞれR、G、Bと呼ぶことにして、手順を説明する。まず、ショットSHkの代表時空間画像の静止領域において、RGB各8階調の512色cのヒストグラムHS(c, SHk)を求める。同様に、動領域のヒストグラムHM(c, SHk)を求める。次に、画素数が少ない色をノイズとみなして除外するために、構成色の最低画素数 θH を設定し、次式を用いて静止構成色HSV_k、動構成色HMT_k (kは1以上の整数) を求める。

【0118】

+mの間で共通する静止構成色、動構成色を、それぞれ静止共通色HSCtop, m、動共通色HMT_{top, m} (top, mは1以上の整数) と呼ぶこととし、次式を用いて計算する。

【0120】

【数4】

Ctop, mをもつ画素がショットSHkの代表時空間画像中に

占める割合」を画像共通色比率AMC (k,top,m)と呼ぶこととし、次式を用いて計算する。

【0121】

【数5】

$$AMC(k, top, m) = \frac{\sum_{c \in HSC_{top, m}} HS(c, SHk) + \sum_{c \in HMC_{top, m}} HM(c, SHk)}{\sum_{c \in HSV_k} HS(c, SHk) + \sum_{c \in HMV_k} HM(c, SHk)} \quad \dots\dots (17)$$

手順1606では、共通色比率条件の成立を表す式

$$AMC(k, top, m) \geq \theta_{shot}, top \leq k \leq top+m \quad \dots\dots (18)$$

が、すべてのkに対して成立するとき、ショットSHtop～SHtop+mが共通色比率条件を満足するとみなす。どこまでのショットが共通色比率条件を満たすか調べるために、すべてのkに対して(18)式が成立する場合には、mに1を足してから手順1604に戻る。そうでない場合には、手順1607に進む。

【0122】本実施例では、時系列の2つのショットの間の画像共通色比率を用いてショット群の末尾を求める。まず、手順1607で、(17)式で定義した画像

$$SS(k) = \min(AMC(k, k, 1), AMC(k+1, k, 1)) \quad \dots\dots (19)$$

を計算し、ショットSHk、SHk+1の統合のための優先度とみなす。

【0124】手順1609では、共通色比率条件を満たすショットSHtop～SHtop+m-1と、その次のショットSHto

$$SS(kmin) \leq SS(k), top \leq k \leq top+m-1 \quad \dots\dots (20)$$

を満たすSS(kmin)を求め、図21(a)及び(b)の例のように、優先度が最小になるショットSHkminをショット群の末尾とみなす。(図21(a)及び(b)では2番目のショット)そして、以上に述べてきた手順1610を終了する。

【0125】なお、ショット群を求める方法は各種考えられる。例えば、時系列のショットで、時間長がほぼ等しい場合には、内容が類似するので、一つのショット群とみなすことができる。したがって、本実施例と異なる方法を用いて、ショット群を求めてもよい。

【0126】また、手順1606の後、手順1607に進まずに、共通色比率条件を満たすショットの末尾SHk+m-1をショット群の末尾とみなして、手順410を終了してもよい。

【0127】図4の手順410の終了後、手順411に進む。手順411では、各ショット群から一つずつの部分動画像を選択する。ただし、部分動画像の選択方法は、ショット群の先頭ショットの先頭部分、ショット群の内容を代表するショットの先頭部分など、各種の方法がある。また、部分動画像の時間長の設定方法についても、固定長にする方法、音声の内容によって時間長を変化させる方法など、各種の方法が考えられる。

共通色比率AMC(k,k,1)、AMC(k+1,k,1)(kはtop以上top+m以下の整数)を計算する。前者AMC(k,k,1)は「2つのショットSHk、SHk+1の静止共通色と動共通色が、前のショットSHkの代表時空間画像中に占める割合」を表す。後者AMC(k+1,k,1)は、「2つのショットSHk、SHk+1の静止共通色と動共通色が、後のショットSHk+1の代表時空間画像中に占める割合」を表す。

【0123】手順1608では、画像共通色比率の最小値

p+mの中から、ショット群の末尾を求める。これらのショットの間の統合の優先度SS(top)～SS(top+m-1)の中から、優先度が最小であることを表す式

【0128】なお、手順411では、部分動画像を選択したが、各ショット群から一つずつの静止画像を選択してもよい。この場合、選択した静止画像を次々に表示すると、要約映像になる。また、選択した静止画像を縮小して一覧表示をしてもよい。一覧表示結果は、要約映像と同等に扱うことができる。

【0129】手順411の終了後、手順412に進む。手順412は、ショット群の特徴を明示するための処理であり、手順411で選択した部分動画像を加工処理する。

【0130】本実施例では、ショット群に含まれるショットが共通色比率条件を満たす。すなわち、ショット群を構成する各ショットに、手順1607で求めた静止共通色、動共通色が存在する。

【0131】静止共通色、動共通色は、ショット群に共通して存在する物体の色を表すので、ショット群の特徴とみなすことができる。例えば、類似背景のショットならば背景の色を表し、類似被写体のショットならば、被写体の色を表す。そこで、手順412では、静止共通色と動共通色の画素を通常と同様に表示し、これ以外の画素を、半分の輝度で表示する。なお、静止共通色と動共通色を強調するものならば、静止共通色と動共通色を除

いた画素の輝度をゼロにするなど、どのような加工処理方法を用いてもよい。

【0132】また、画像の加工処理方法は、ショットの統合方法にあわせて決めるものであり、様々なバリエーションが存在する。例えば、時間長がほぼ等しいショットを統合する場合には、図22のように、部分動画像の下側にショットの平均時間長を表示する加工処理方法などが考えられる。

【0133】手順412の終了後、手順402に戻る。なお、ショット群の特徴を明示する必要がない場合には、手順412を実行せずに、手順411の終了後に手順402に戻ってもよい。

【0134】以上が、図1における第1の計算機5の映像要約処理の一実施例であり、図2の手順201の詳細な説明である。

【0135】映像要約処理の終了後、図2の手順202を実行する。手順202では、図1における第1の計算機5がビデオディスク装置1やVTR2を制御して、処理対象映像と手順201で要約した映像を再生する。再生された映像は図1の映像圧縮装置4においてMPEG方式で圧縮され、ファイルサーバー6に記憶される。

【0136】ただし、本実施例の手順201は、要約した映像を作成せずに、図3に示したような要約映像の再生方法を作成するので、この要約映像の再生方法にしたがってビデオディスク装置1やVTR2を制御することで、要約映像の再生を実行する。

【0137】なお、必ずしもMPEG方式で圧縮する必要はなく、JPEG方式などの他の圧縮方式で圧縮してもよい。また、要約映像の記憶方法については、すでに述べたように、映像を直接圧縮する方法だけでなく、他の方法が各種考えられる。

【0138】以下では、図2における手順203の具体的な動作である、図1の第2の計算機7の要約映像再生処理について述べる。

【0139】図23は、図1における第2の計算機7の要約映像再生処理の一実施例のフローチャートである。

【0140】手順2301では、映像表示方法の選択を行う。ただし、多数の選択肢が考えられる。例えば、映像の内容を詳しく見たい場合には、「処理対象映像の標準速度再生」を選択すればよい。映像表示時に、ユーザーの見たい映像が普通で再生される。

【0141】また、被写体の動作を中心として内容を把握したい場合には、図4の手順405から手順408までの処理によって作成した「区間変速要約映像」を選択すればよい。映像表示時に、主観的にみて再生速度が許容範囲内に入るようにしながら、映像全体が再生される。

【0142】なお、主観には個人差がある。したがって、ユーザーは、自分の主観に合わせて、再生速度決定方法を選択してもよい。具体的には、次の3つの処理を

実行するかわりに、その一部だけを実行したり、「テクスチャーの細かい映像の部分で、再生速度を相対的に遅くする」処理などを付加してもよい。

【0143】・長時間類似区間で、再生速度を相対的に遅くするための、手順405

・高速動作区間で、再生速度を相対的に遅くするための、手順406

・短時間長ショットで、再生速度を相対的に遅くするための、手順408

さらに、映像に含まれる内容の時間的な流れを知りたい場合には、図4の手順409によって作成した「リズム呈示要約映像」を選択すればよい。映像表示時に、同じ内容の部分をまとめたショットが次々に再生される。

【0144】できるだけ異なる内容の部分を少しづつ見たい場合には、図4の手順410から手順412までの処理によって作成した「部分選択要約映像」を選択すればよい。映像表示時に、時系列の類似する内容のショットをまとめたショット群が次々に再生される。

【0145】図23の手順2302では、手順2301による選択結果にしたがって、必要な映像を図1のファイルサーバー6から呼び出し、その映像を再生する。ただし、要約映像作成のためのフレーム番号情報だけがファイルサーバー6に保存されている場合には、フレーム番号情報にしたがって、処理対象映像を要約しながら再生する。また、ユーザーは、映像再生前に、「どのフレームからどのフレームまで再生するか」を指定してもよいし、「どのフレームから再生するか」だけを指定して、映像を見ながら好みの時点で再生を中断してもよい。

【0146】なお、図4に示した映像要約処理の流れは、手順2301の選択肢にあわせて決めればよい。たとえば、手順2301の選択肢に「リズム呈示要約映像」が含まれない場合には、図4の手順409を実行する必要はない。同様に、「区間変速要約映像」が含まれない場合には、図4の手順405から手順408までを実行する必要はなく、「部分選択要約映像」が含まれない場合には、図4の手順410から手順412までを実行する必要はない。

【0147】また、本実施例では、要約した映像と処理対象映像を図1の第2の計算機7上で再生する例を示したが、ファイルサーバー6と結ばれている他の計算機上で再生してもよい。例えば、図1の第1の計算機5上で再生してもよいし、2台以上の計算機がファイルサーバー6と結ばれている場合には、そのすべての計算機上で再生してもよい。

【0148】以上のように、本実施例の区間変速要約映像によれば、一定時間以上にわたって類似した画像が続く時間区間を検出して長時間類似区間とみなし、長時間類似区間を相対的に速く再生することで、映像が短時間で次々に変化するの、「主観的に再生速度が遅い部

分」が消滅する。また、高速で動く物体を含む時間区間を検出して高速動作区間とみなし、高速動作区間の再生速度を相対的に遅くすることで、高速で動く物体が存在しなくなるので「主観的に再生速度が速い部分」が減少する。さらに、映像を内容ごとにまとめて複数のショットに分割してから、各ショットの表示時間長に下限を設けることで、各内容が一定時間以上表示されるので、「主観的に再生速度が速い部分」が消滅する。このように、主観的にみて、再生速度が許容範囲内に入るようにしながら映像全体を再生することで、従来法に比べてユーザーの疲労が軽減される。また、被写体の動作を中心とした内容を把握することができる。

【0149】また、本実施例のリズム呈示要約映像では、各ショットの表示時間長に上限を設定してから、次の2つの条件の少なくとも一方を満足するように再生速度を決定し、リズムを想起させながら、時系列のショットを次々に再生する。

【0150】リズム条件：ショットの境界と、予め設定した周期のリズムが相関をもつ。

【0151】内容条件：「再生速度決定の基準である速度上限値以下の速度で再生した部分は必ず内容を把握できる」ということを保証するように速度上限値が設定されているとき、各ショットの少なくとも一部分の再生速度が速度上限値以下になる。

【0152】リズム条件を満足するように映像を要約することで、ショットの変化を予測できるので、すべてのショットを見落とさずに見ることができる。

【0153】内容条件を満足するように映像を要約することで、ユーザーは、すべてのショットにおいて、少なくともその一部分の内容を把握することができる。同じ内容の部分をまとめたものがショットになっているので、すべてのショットの内容を把握できることになる。

【0154】ただし、リズム条件を満足しない場合には、ショットの変化を予測できないので、ショットの見落としが発生する。内容条件を満足しない場合には、一部のショットの内容を把握できない。リズム条件と内容条件を同時に満足することが望ましい。

【0155】さらに、本実施例の部分選択要約映像では、時系列のショットの間の相関を用いて、類似ショットを統合してショット群を作成してから、各ショット群から選択した部分動画像を次々に再生する。この要約映像を用いることで、できるだけ異なる内容の部分を効率よく見ることができる。

【0156】

【発明の効果】以上のように本発明は、第1に、一定時間以上にわたって類似した画像が続く時間区間を検出して長時間類似区間とみなし、長時間類似区間を相対的に速く再生することで、映像を短時間で次々に変化させ、「主観的に再生速度が遅い部分」を消滅させることができる。また、高速で動く物体を含む画像の時間区間を検

出して高速動作区間とみなし、高速動作区間の再生速度を相対的に遅くすることで、高速で動く物体を消滅させ、「主観的に再生速度が速い部分」を減少させることができる。さらに、同じ内容の部分をショットにまとめてから、各ショットの表示時間長に下限を設けることで、各内容を一定時間以上表示させ、「主観的に再生速度が速い部分」を消滅させることができる。このように、主観的にみて、再生速度が許容範囲内に入るようにしながら映像全体を再生することで、従来法に比べてユーザーの疲労が軽減される。また、被写体の動作を中心とした内容を把握することができる。

【0157】また第2に、各ショットの表示時間長に上限を設定してから、次の2つの条件の少なくとも一方を満足するように再生速度を決定し、リズムを想起させながら、時系列のショットを次々に再生する。

【0158】リズム条件：ショットの境界と、予め設定した周期のリズムが相関をもつ。

【0159】内容条件：「再生速度決定の基準である速度上限値以下の速度で再生した部分は必ず内容を把握できる」ということを保証するように速度上限値が設定されているとき、各ショットの少なくとも一部分の再生速度が速度上限値以下になる。

【0160】リズム条件を満足するように映像を要約することで、ショットの変化を予測できるので、すべてのショットを見落とさずに見ることができる。

【0161】内容条件を満足するように映像を要約することで、ユーザーは、すべてのショットにおいて、少なくともその一部分の内容を把握することができる。同じ内容の部分をまとめたものがショットになっているので、すべてのショットの内容を把握できることになる。

【0162】ただし、リズム条件を満足しない場合には、ショットの変化を予測できないので、ショットの見落としが発生する。内容条件を満足しない場合には、一部のショットの内容を把握できない。リズム条件と内容条件を同時に満足することが望ましい。

【0163】さらに第3に、時系列のショットの間の相関を用いて、類似ショットを統合してショット群を作成してから、各ショット群から選択した部分動画像を次々に再生する。この要約映像を用いることで、できるだけ異なる内容の部分を効率よく見ることができる。

【図面の簡単な説明】

【図1】本発明の一実施例における映像要約装置の全体システム図

【図2】同実施例における映像要約装置の動作のフローチャート

【図3】同実施例におけるファイル形式による要約映像の表現の概念図

【図4】同実施例における映像要約処理のフローチャート

【図5】同実施例における長時間類似区間検出処理のフ

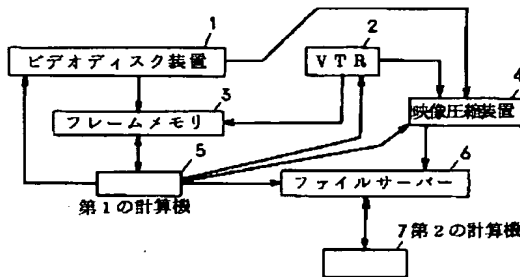
ローチャート

- 【図6】同実施例における部分領域の作成例を示す図
 【図7】同実施例における映像のゆるやかな映像変化をする部分を示す図
 【図8】同実施例における高速動作区間検出処理のフローチャート
 【図9】同実施例における画面上の物体が高速に移動する場合を示す図
 【図10】同実施例における(6)式のみを用いた画素の比較を示す図
 【図11】同実施例における(6)、(7)式における、位置 p の画素の比較を示す図
 【図12】同実施例における高速動作画素検出を実行する領域を示す図
 【図13】同実施例におけるリズム呈示要約処理のフローチャート
 【図14】同実施例におけるリズム条件を満たす再生方法の概念図
 【図15】同実施例におけるリズム呈示要約映像の再生方法の概念図
 【図16】同実施例におけるショット統合処理のフロー

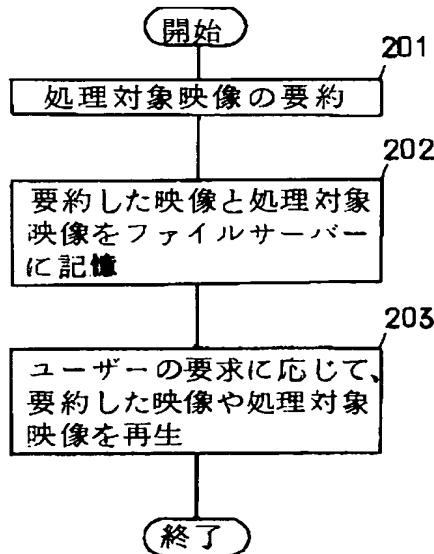
チャート

- 【図17】同実施例における類似背景のショットの概念図
 【図18】同実施例における類似被写体のショットの概念図
 【図19】同実施例における動領域の概念図
 【図20】同実施例における動領域の概念図
 【図21】同実施例における末尾決定処理の概念図
 【図22】同実施例における部分動画像の加工方法の概念図
 【図23】同実施例における要約映像再生処理のフローチャート
 【符号の説明】
 1 ビデオディスク装置
 2 VTR
 3 フレームメモリ
 4 映像圧縮装置
 5 計算機
 6 ファイルサーバー
 7 計算機

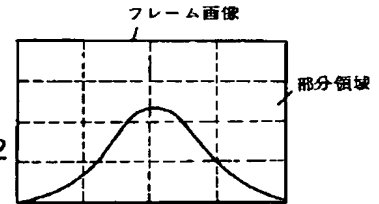
【図1】



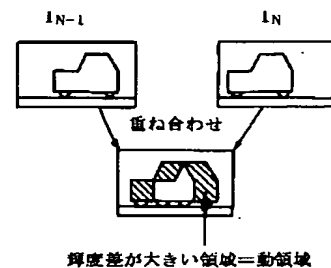
【図2】



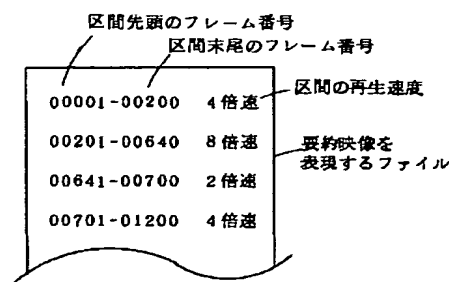
【図6】



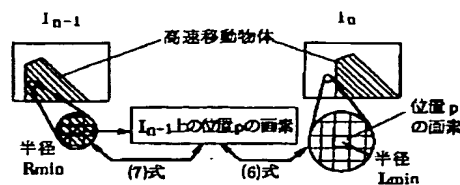
【図20】



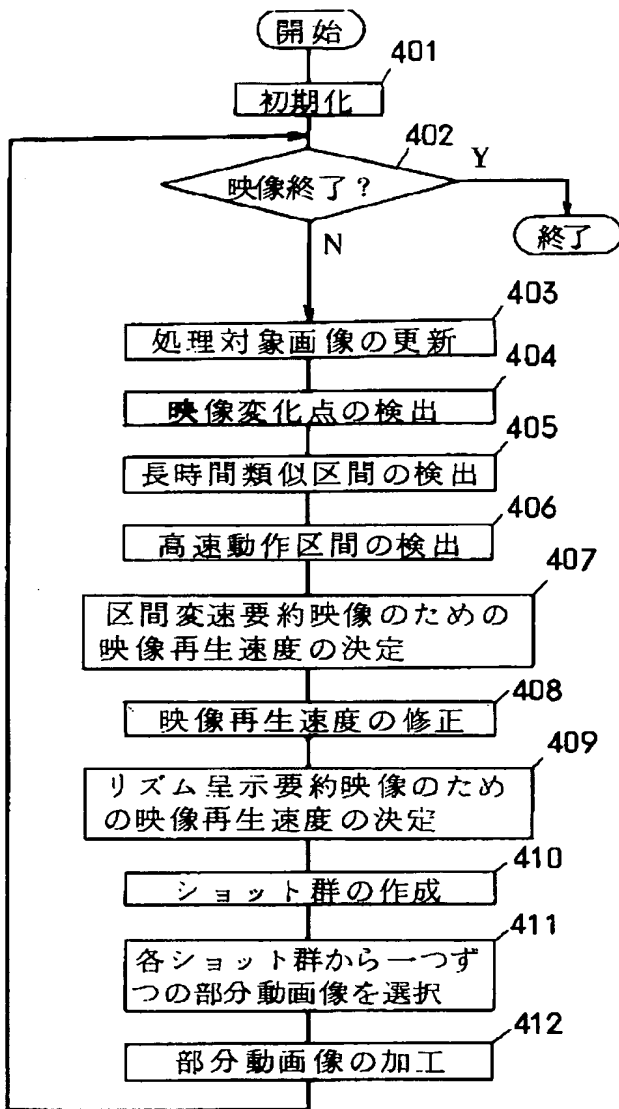
【図3】



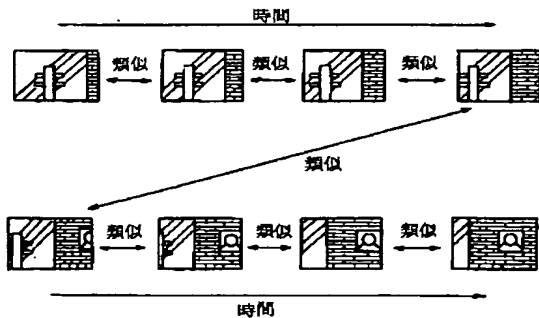
【図9】



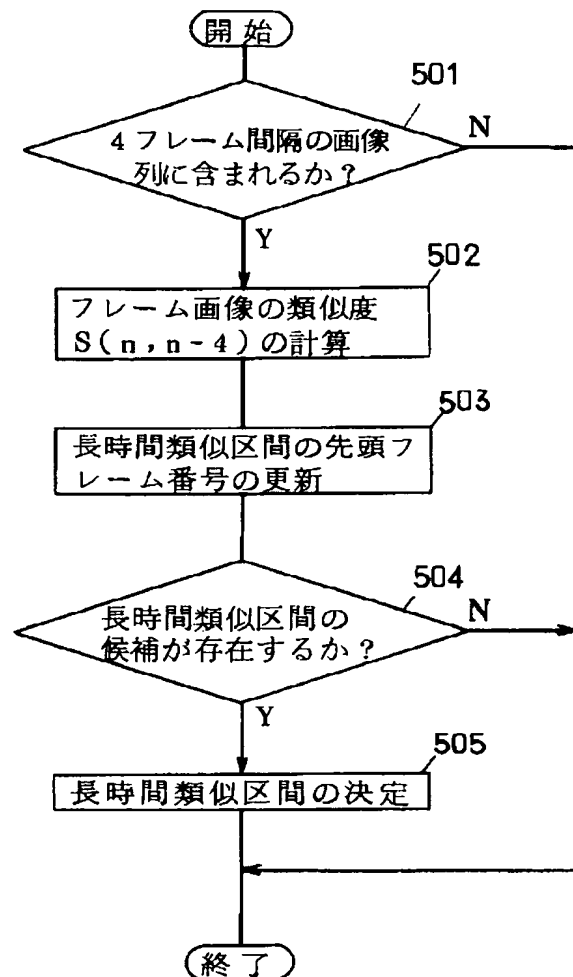
【図4】



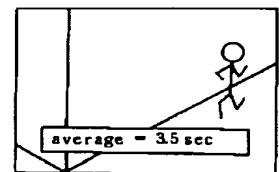
【図7】



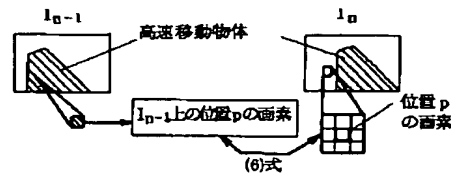
【図5】



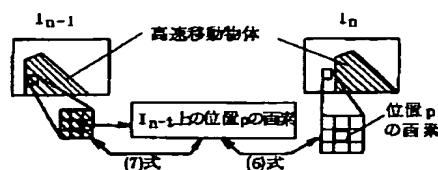
【図22】



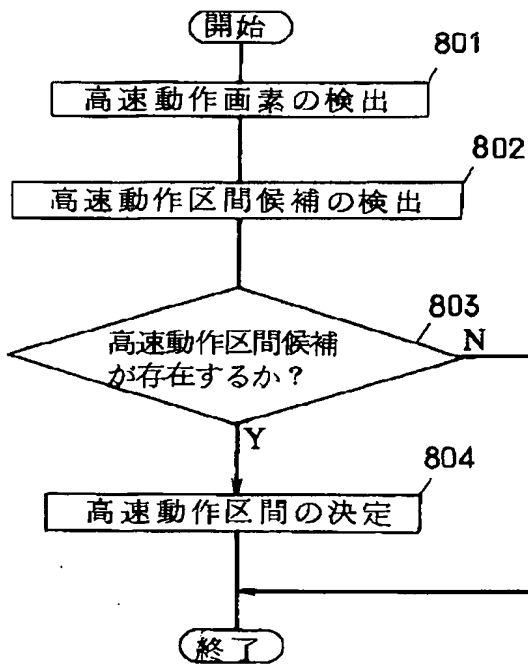
【図10】



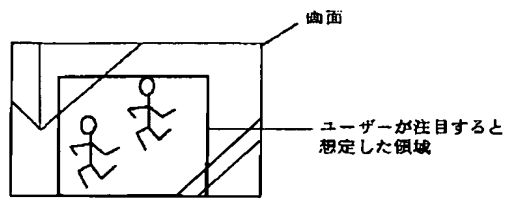
【図11】



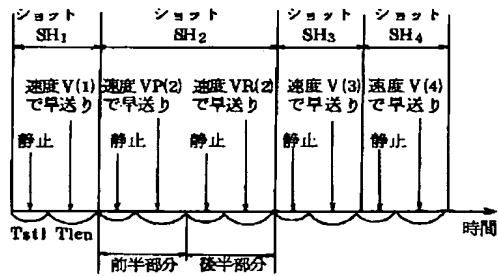
【図8】



【図12】

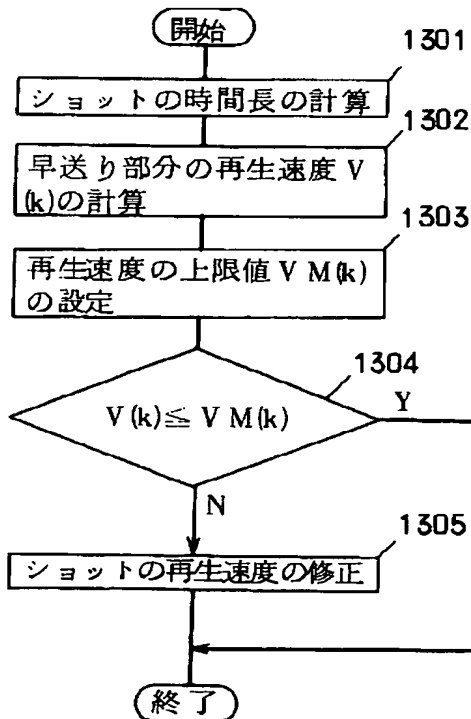


【図15】

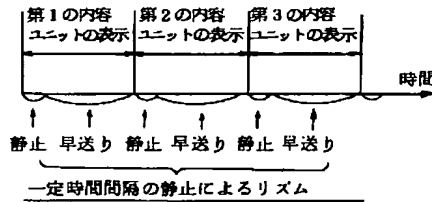


【図19】

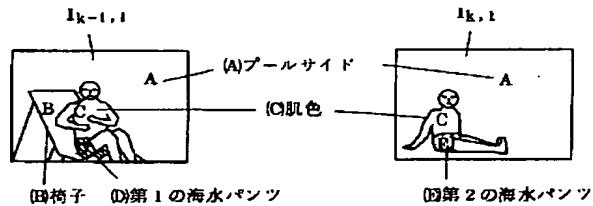
【図13】



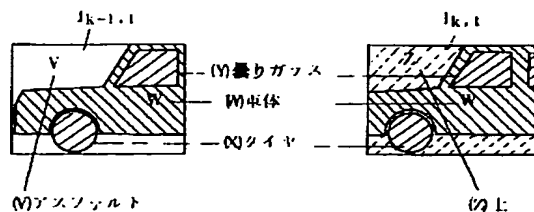
【図14】



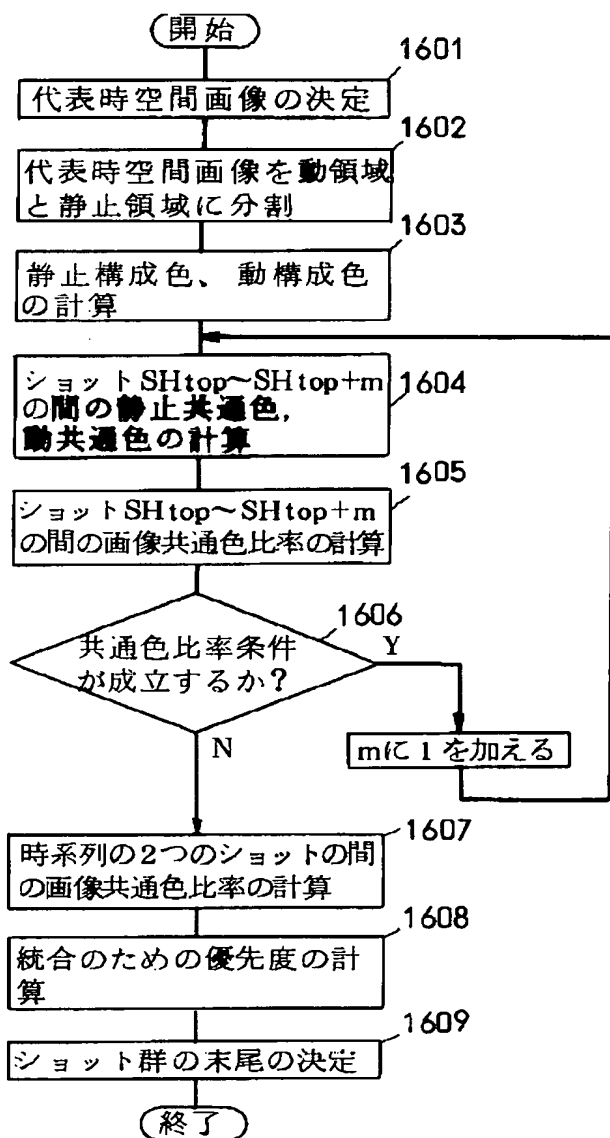
【図17】



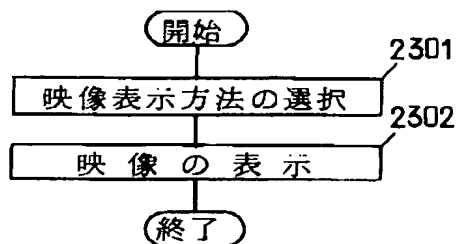
【図18】



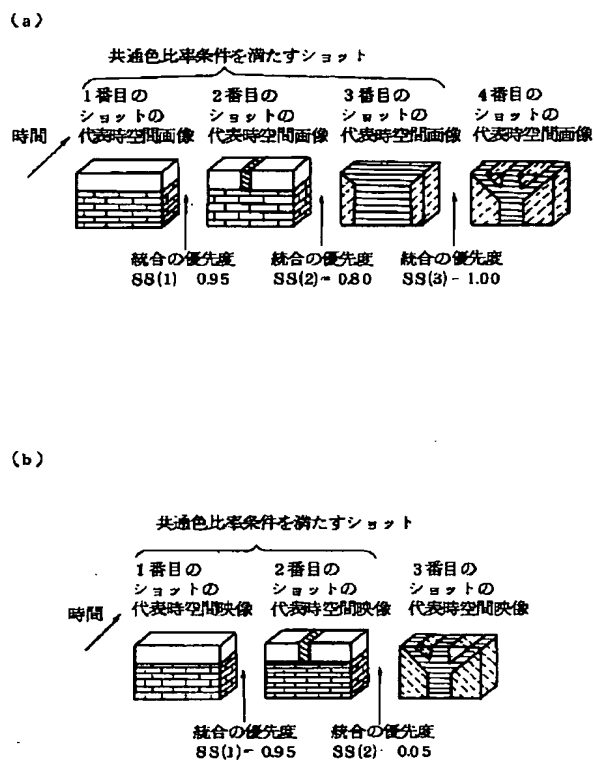
【図16】



【図23】



【図21】



PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-251540

(43)Date of publication of application : 27.09.1996

(51)Int.Cl. H04N 5/93

(21)Application number : 07-046970 (71)Applicant : MATSUSHITA ELECTRIC IND
CO LTD

(22)Date of filing : 07.03.1995 (72)Inventor : YAMADA SHIN
KANAMORI KATSUHIRO

(54) VIDEO SUMMARIZING METHOD

(57)Abstract:

PURPOSE: To summarize automatically video by solving a problem that a desired content cannot be grasped when the video summarized by the user is viewed so as to provide efficiently a content among the operation of an object with respect to the method summarizing and displaying the video to be used for a device supporting retrieval, editing, processing and glance of the video.

CONSTITUTION: The frame image of a video signal received by a video disk device 1 or a VTR 2 is received by a frame memory 3, a 1st computer 5 gives a control signal to the frame memory 3 to receive a frame image and the video is summarized by processing a time series frame image and date are stored in a file server 6. A 2nd computer 7 calls a summarized video from the file server 6 on request by the user and reproduces the video.

LEGAL STATUS [Date of request for examination] 09.03.1999
[Date of sending the examiner's decision of rejection] 12.06.2001
[Kind of final disposal of application other than the examiner's decision of rejection or
application converted registration]
[Date of final disposal for application]
[Patent number] 3240871
[Date of registration] 19.10.2001
[Number of appeal against examiner's decision of rejection] 2001-012113
[Date of requesting appeal against examiner's decision of rejection] 12.07.2001
[Date of extinction of right]

*** NOTICES ***

JPO and INPIT are not responsible for any
damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect
the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] The image epitome approach which calculates the similarity between the
frame images of the image used as the candidate for an epitome, detects the long
duration similar section which is the time amount section when an image similar over
beyond the time amount on which it decided beforehand continues, and reproduces
said long duration similar section quickly compared with the time amount sections
other than said long duration similar section.

[Claim 2] The image epitome approach according to claim 1 which divides two frame
images into two or more subregions, respectively, and calculates the similarity of said
two frame images using the similarity of said subregion.

[Claim 3] The image epitome approach which detects the high-speed operation
section which is the time amount section which the body on a screen moves quickly
compared with other time amount sections, and reproduces said high-speed operation

section late from the frame image of the image used as the candidate for an epitome compared with the time amount sections other than said high-speed operation section.

[Claim 4] The image epitome approach according to claim 3 of detecting the high-speed operation section in the field in which the frame image was decided beforehand.

[Claim 5] The frame image IM of time series (M is the one or more natural numbers) is sampled from the image used as the candidate for an epitome. The pixel PN in the specific frame image IN (N is the natural number of 1 [or more] and under M) The brightness difference between the pixels around [corresponding to said pixel PN] pixel PN+1 is calculated in the following frame image IN+1. The image epitome approach according to claim 3 or 4 of asking for said pixel PN to which said brightness difference exceeds the threshold θ_{W1} set up beforehand per each frame image, and asking for the high-speed operation section using said pixel PN.

[Claim 6] The frame image IM of time series (M is the one or more natural numbers) is sampled from the image used as the candidate for an epitome. The pixel PN in the specific frame image IN (N is the natural number of 1 [or more] and under M) The 1st brightness difference which is a brightness difference between the pixels around said pixel PN is calculated. Said pixel PN The 2nd brightness difference which is a brightness difference between the pixels around [corresponding to said pixel PN] pixel PN+1 is calculated in the following frame image IN+1. Per each frame image The image epitome approach according to claim 3 or 4 of asking for said pixel PN to which said 1st brightness difference becomes less than θ_{W1} threshold set up beforehand, and said 2nd brightness difference exceeds said threshold θ_{W1} , and asking for the high-speed operation section using said pixel PN.

[Claim 7] The image epitome approach according to claim 3 or 4 of asking for the high-speed operation section using a motion vector.

[Claim 8] Claim 1 which packs the frame image of the image used as the candidate for an epitome for every contents, divides into two or more shots, and prepares a minimum in the display time length of said shot according to the contents of said shot thru/or claim 7 are the image epitome approach of a publication either.

[Claim 9] The image epitome approach which carries out sequential playback of said shot while pack the frame image of the image used as the candidate for an epitome for every contents, and it divides into two or more shots, determining reproduction speed so that the boundary of said shot and the rhythm of the period set up beforehand may have correlation, and reminding a rhythm.

[Claim 10] The image epitome approach according to claim 9 of reminding a rhythm using a sound.

[Claim 11] between the time amount Tstl which set up beforehand the frame image which interrupts playback at intervals of the time amount Tlen set up beforehand, and

is displayed at the time -- displaying ($T_{len}+T_{stl}$) -- the image epitome approach according to claim 9 of determining reproduction speed so that the rhythm made into a period may be generated and said rhythm may have the boundary of a shot, and correlation.

[Claim 12] The image epitome approach which packs the frame image of the image used as the candidate for an epitome for every contents, divides into two or more shots, calculates the rate upper limit which is the criteria of reproduction-speed decision according to the contents of the image, reproduces some [at least] shots at the rate of below said rate upper limit, determines the reproduction speed of said shot so that it may become below the time-amount length upper limit that is a threshold which the display-time length of said shot set up beforehand, and carries out the sequential playback of the shot.

[Claim 13] Pack the frame image of the image used as the candidate for an epitome for every contents, and it divides into two or more shots. According to the contents of said image, the rate upper limit which is the criteria of reproduction speed decision is calculated. When reproduction speed required in order to display said whole shot by the conventional-time length $Trhy$ who set up beforehand becomes said below rate upper limit The reproduction speed of said shot is determined so that the display time length of said shot may turn into said conventional-time length $Trhy$. When reproduction speed required in order to display said whole shot by said conventional-time length $Trhy$ exceeds said rate upper limit The image epitome approach which sets up the display time length of said shot the twice of said conventional-time length $Trhy$, determines the reproduction speed of said shot so that the reproduction speed between $Trh(ies)$ in the first half of said display time length may become said below rate upper limit, and carries out sequential playback of the shot of time series.

[Claim 14] The image epitome approach according to claim 12 or 13 of setting a rate upper limit as a large value compared with the shot which does not include said long duration similar section at the shot which calculates the similarity between frame images, detects the long duration similar section which is the time amount section when contents similar over beyond the time amount on which it decided beforehand continue, and includes said long duration similar section.

[Claim 15] The image epitome approach according to claim 14 which divides two frame images into two or more subregions, respectively, and calculates the similarity of two frame images using the similarity of said subregion.

[Claim 16] The image epitome approach according to claim 12 or 13 of setting a rate upper limit as a small value compared with the shot which does not include said high-speed operation section at the shot which detects the high-speed operation section which is the time amount section which the body on a screen moves quickly compared with other time amount sections, and includes said high-speed operation

section from the frame image of the image used as the candidate for an epitome.

[Claim 17] The image epitome approach according to claim 16 of detecting the high-speed operation section in the field in which the frame image was decided beforehand.

[Claim 18] The frame image IM of time series (M is the one or more natural numbers) is sampled from the image used as the candidate for an epitome. The pixel PN in the specific frame image IN (N is the natural number of 1 [or more] and under M) The brightness difference between the pixels around [corresponding to said pixel PN] pixel PN+1 is calculated in the following frame image IN+1. The image epitome approach according to claim 16 or 17 of asking for said pixel PN to which said brightness difference exceeds the threshold $\theta W1$ set up beforehand per each frame image, and asking for the high-speed operation section using said pixel PN.

[Claim 19] The frame image IM of time series (M is the one or more natural numbers) is sampled from the image used as the candidate for an epitome. The pixel PN in the specific frame image IN (N is the natural number of 1 [or more] and under M) The 1st brightness difference which is a brightness difference between the pixels around said pixel PN is calculated. Said pixel PN The 2nd brightness difference which is a brightness difference between the pixels around [corresponding to said pixel PN] pixel PN+1 is calculated in the following frame image IN+1. Per each frame image The image epitome approach according to claim 16 or 17 of asking for said pixel PN to which said 1st brightness difference becomes less than $\theta W1$ threshold set up beforehand, and said 2nd brightness difference exceeds said threshold $\theta W1$, and asking for the high-speed operation section using said pixel PN.

[Claim 20] The image epitome approach according to claim 16 or 17 of asking for the high-speed operation section using a motion vector.

[Claim 21] The image epitome approach which divides the frame image of the image used as the candidate for an epitome into two or more shots, unifies said shot similar using correlation between said shots of time series, regards it as a shot group, chooses from said shot group the partial dynamic image which is a frame image of two or more time series, and carries out sequential playback of said partial dynamic image.

[Claim 22] The image epitome approach according to claim 21 which searches for the description which is common from all the shots contained in a shot group, processes a partial dynamic image based on said description, and carries out sequential playback of the partial dynamic image after processing.

[Claim 23] The image epitome approach according to claim 22 which searches for the description which is common from all the shots contained in a shot group, processes a partial dynamic image based on said description, reduces the frame image of the head of the partial dynamic image after processing, and indicates by list on a screen.

[Claim 24] Claim 21 which detects the shot of the time series by which the similar background and the similar photographic subject are photoed, and is made into a shot

group thru/or claim 23 are the image epitome approach of a publication either.

[Claim 25] Claim 21 which processes the image between representation space-time which is a frame image chosen from shots, divides into the motion area which is a field influenced of a change according the image between said representation space-time to a motion, and the quiescence field which are fields other than said motion area, creates the histogram of a color for every field, and searches for correlation between shots using the histogram of said color thru/or claim 24 are the image epitome approach of a publication either.

[Claim 26] The image epitome approach according to claim 25 of processing a partial dynamic image in a shot group using the intersection of ***** which is a color which appears in a motion area, and the intersection of the quiescence configuration color which is a color which appears in a quiescence field.

[Claim 27] Claim 21 which time amount length detects the shot which continues in time almost equally, and considers is a shot group thru/or claim 23 are the image epitome approach of a publication either.

[Claim 28] Claim 8 which considers that the part continuously photoed in time with the video camera is a shot thru/or claim 27 are the image epitome approach of a publication either.

[Claim 29] Claim 8 which considers that one scene of a scenario is a shot thru/or claim 27 are the image epitome approach of a publication either.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention summarizes the image which started the approach of supporting retrieval of an image, edit, processing, Hayami, etc., especially was stored in the video tape or the videodisk, and relates to the image epitome approach which gives playback or an indication.

[0002]

[Description of the Prior Art] In recent years, a computer etc. is applied and the research on the approach of supporting retrieval of an image, edit, processing, Hayami, etc. prospers. As the example, in order to grasp the contents of the image of video and a movie for a short time, there is the image epitome approach which reproduces the whole part and whole image of an image for a short time.

[0003] As a conventional approach, the rushes reclaiming process is proposed the adjustable-speed reclaiming process (an Otsuji, Tomura, and "subjectivity evaluation of dynamic-image high-speed browsing" 1993 Institute of Electronics, Information and

Communication Engineers spring convention, SD-9-3), and the whole shot the rapid-traverse reclaiming process, the contents compression display process approach of an image given in JP,4-237284,A, and the whole shot. In addition, a shot is the unit of the image often used in fields, such as image edit, and it is close to the smallest unit of the contents of an image.

[0004] The conventional rapid-traverse reclaiming process is the image epitome approach which thins out a frame image in a fixed time interval, and is reproduced.

[0005] The contents compression method of presentation of an image is the image epitome approach that judge the display significance of each frame image and a frame image with a higher significance spares much time amount for a display using the variation between the frame images of time series. For utilization, the part which judges the display significance of a frame image is important.

[0006] The brightness variation of the pixel unit between contiguity frame images is sensitive to a motion of the body in a screen. Then, few places of a motion are flown and the method of considering that the brightness variation of this pixel unit is the above-mentioned display significance is proposed as the display significance judging approach used to see slowly the place with many motions.

[0007] On the other hand, like [the brightness variation of the frame unit between contiguity frame images is comparatively insensible to migration of the body within a shot, and / in case a shot changes], when the luminance distribution inclination of the whole frame changes, a big value comes out. Then, the method of considering that the brightness variation of this frame unit is the above-mentioned display significance is proposed as the display significance judging approach used to see change of a shot carefully.

[0008] An adjustable-speed reclaiming process is the image epitome approach which reproduces an image the conventional whole shot, controlling reproduction speed to make display time of each shot regularity.

[0009] Moreover, a rushes reclaiming process is the image epitome approach which reproduces the head part of a shot one after another by standard speed the whole shot.

[0010] Finally, it supplements with explanation about the shot. The part continuously photoed in time with one video camera is called a shot. A shot is close to the smallest unit of the contents of an image, as mentioned above. Moreover, the part connected by edit and the part which interrupted photography of a video camera serve as "change of a shot."

[0011] However, it may be considered as an exception that the part from which the contents of an image change with camera actuation of a pan, a zoom, etc. is "change of a shot." In this case, compared with the case where an exception is not taken into consideration, a shot becomes close to the smallest unit of the contents of an image.

[0012] As an approach of dividing an image into a shot automatically, the image

change model method (Yamada, Fujioka, Kanamori, Matsushima, Sakauchi, the "scene change detection approach of an image including the edit effectiveness", multimedia, and image processing symposium'94) etc. is proposed.

[0013]

[Problem(s) to be Solved by the Invention] However, in the above-mentioned rapid-traverse reclaiming process, since it reproduced with constant speed regardless of the contents of an image, it saw on "subjectivity target and a part with quick reproduction speed and late partial" existed, contents grasp was difficult and the user had a technical problem called the fatigue and a cone.

[0014] Moreover, in the above-mentioned contents compression method of presentation of an image, only the approach of using the brightness variation of the pixel unit between contiguity frame images, and the approach using the brightness variation of a frame unit were proposed as the judgment approach of display significance. Since the brightness variation of a pixel unit was unsuitable to subjective motion evaluation when the former is used, it saw on "subjectivity target and a part with quick reproduction speed and late partial" existed, grasp of the contents was difficult and the user had a technical problem called the fatigue and a cone. Moreover, since only the head frame image of each shot was displayed when the latter is used, it had the technical problem that it could not use to grasp the contents focusing on actuation of a photographic subject.

[0015] The above-mentioned whole shot, with the adjustable-speed reclaiming process, since reproduction speed was decided by the time amount length of a shot, a user has not grasped the contents of some shots but it had the technical problem that it could not use to know the time flow of the contents included in an image.

[0016] It becomes important [it is important to grasp the contents of each shot which is the unit of the contents, and], in order to avoid an oversight of a shot that change of a shot can be predicted to know the time flow of the contents. However, the whole shot, with an adjustable-speed reclaiming process, if the case "where a shot changes in the part connected by edit and the part which interrupted photography of a video camera" is removed, change of a shot cannot be predicted. In the time of considering that the part photoed by the same camera actuation is the above-mentioned shot, and treating it, and the time of considering that one scene of a scenario is the above-mentioned shot, and treating it, change of a shot could not be predicted but it had the technical problem that an oversight of a shot occurred.

[0017] With an image, the combination of two or more shots serves as a unit of contents higher order than a shot like a scene. The above-mentioned whole shot, with the adjustable-speed reclaiming process, since all the shots were reproduced little by little, the part of similar contents will be reproduced continuously and it had the technical problem that it could not use to see efficiently the part of different contents as much as possible.

[0018] This invention solves the technical problem of the above-mentioned conventional technique, and it aims at offering the image epitome approach that a user can grasp the desired contents.

[0019]

[Means for Solving the Problem] In order to attain this purpose, the time amount section of the image containing the body which moves at high speed is detected, and it is regarded [1st] as the high-speed operation section. Moreover, the time amount section when an image similar over beyond fixed time amount continues is detected, and it is regarded as the long duration similar section. And the high-speed operation section is reproduced at a low speed compared with the time amount sections other than the high-speed operation section, and the long duration similar section is reproduced at high speed compared with the time amount sections other than the long duration similar section. Furthermore, a minimum is prepared for the display time length of each shot, after summarizing an image for every contents and dividing into two or more shots.

[0020] The shot of time series is reproduced one after another, reminding the rhythm of the period set up beforehand, after summarizing an image for every contents and dividing [2nd] into a shot. However, an upper limit is prepared for the display time length of each shot. Moreover, it is determined that reproduction speed will satisfy at least one side of the following two conditions.

[0021] Rhythm conditions: The boundary of a shot and the rhythm of the period set up beforehand have correlation.

[0022] Contents conditions : When the rate upper limit is set up so that it may guarantee "the part reproduced at the rate of below the rate upper limit which is the criteria of reproduction speed decision can surely grasp the contents", a part of [at least] reproduction speed of each shot becomes below a rate upper limit.

[0023] The image for an epitome is divided into the 3rd at a shot. After it, using correlation between the shots of time series, a similar shot is unified and a shot group is created. And every one partial dynamic image is chosen from each shot group, and a partial dynamic image is reproduced one after another.

[0024]

[Function] Since the body which moves at high speed stops existing and each contents are displayed [1st] beyond fixed time amount by these configurations, "it is a part with quick reproduction speed subjectively" disappears. Moreover, since an image changes one after another for a short time, "it is a part with slow reproduction speed subjectively" disappears. Therefore, fatigue of a user is mitigated compared with a conventional method. Moreover, since this approach reproduces the whole image, seeing subjectively and making it reproduction speed enter in tolerance, it can grasp contents centering on actuation of a photographic subject.

[0025] To the 2nd, the part of the same contents is summarized to a shot and each is

reproduced for a short time. Therefore, since change of a shot can be predicted when summarizing an image as satisfying rhythm conditions, it can see, without overlooking no shots.

[0026] On the other hand, when summarizing an image as satisfying contents conditions, a user can grasp some of the contents at least in all shots. Since it is the shot which summarized the part of the same contents, the contents of all the shots can be grasped.

[0027] In addition, since the part of the same contents is summarized, a shot is created and these are reproduced one after another, this means is used to know the time flow of the contents included in an image. However, since change of a shot cannot be predicted when not satisfying rhythm conditions, an oversight of a shot occurs. When not satisfying contents conditions, the contents of some shots cannot be grasped. It is desirable that it is satisfied with coincidence of rhythm conditions and contents conditions.

[0028] Since the part of the contents to which time series is similar is summarized, a shot group is created and these parts are reproduced [3rd] one after another, the part of different contents as much as possible can be seen efficiently.

[0029]

[Example] Hereafter, one example of this invention is explained, referring to a drawing.

[0030] Drawing 1 is the whole image epitome equipment system chart in one example of this invention. In drawing 1 , 1 and 2 are the input devices of the image (it is hereafter called a processing-object image) used as a processing object, 1 is videodisk equipment and 2 is VTR. Moreover, 3 is a frame memory which captures the frame image of the video signal outputted from videodisk equipment 1 or VTR2. 4 is an image compression equipment which compresses the video signal outputted from videodisk equipment 1 or VTR2. 5 is the 1st calculating machine and controls videodisk equipment 1, VTR2, a frame memory 3, and the image compression equipment 4. Moreover, a control signal is sent to a frame memory 3, a frame image is captured, and an image is summarized by processing the frame image of time series. 6 is a file server which memorizes the data sent from the image compressed with the image compression equipment 4, or the 1st calculating machine 5, and a frame image. 7 is the 2nd computer which calls an epitome image and a processing-object image from a file server 6, and reproduces the image according to a demand of a user.

[0031] About the image epitome equipment constituted as mentioned above, actuation of the whole is explained using the flow chart shown in drawing 2 .

[0032] In a procedure 201, an image is summarized because the 1st calculating machine 5 in drawing 1 processes the frame image of time series, controlling videodisk equipment 1, VTR2, and a frame memory 3.

[0033] In a procedure 202, controlling videodisk equipment 1, VTR2, and the image compression equipment 4, the 1st calculating machine 5 compresses the image and

processing-object image which were summarized, and memorizes to a file server 6.

[0034] In addition, not only the approach of compressing an image directly about the store method of an epitome image but other approaches are various idea ****. For example, as shown in drawing 3 , when an epitome image can be expressed by file format using the frame number of a processing-object image, the information on a frame number may be memorized to a file server 6 instead of an epitome image. Moreover, controlling videodisk equipment 1, VTR2, and a frame memory 3, the head frame image of each shot group mentioned later may be captured, and you may memorize to a file server 6 instead of an epitome image. However, it is equivalent to showing the epitome of an image to reduce the memorized head frame image in this case, and to indicate by list.

[0035] In a procedure 203, according to a demand of a user, the 2nd calculating machine 7 calls an epitome image and a processing-object image from a file server 6, and reproduces the image.

[0036] Below, the image epitome processing of the 1st computer 5 which is concrete actuation of the procedure 201 in drawing 2 is explained.

[0037] Drawing 4 is the flow chart of one example of image epitome processing of the 1st calculating machine 5 in drawing 1 .

[0038] First, in a procedure 401, initialization of various parameters used for image epitome processing and a setup of a threshold are performed. Moreover, when calculating characteristic quantity, such as "the similarity between frame images", a high-speed operation pixel, a quiescence configuration color, *****, a quiescence common color, and a ** common color, in the procedure mentioned later, in order to use, the processing-object frame image In (n is the one or more natural numbers) is captured from a frame memory 3, and this frame image is memorized.

[0039] Next, in a procedure 402, when it judges whether the image used as a processing object was completed and an image is completed, image epitome processing is ended. When the image is not completed, it progresses to a procedure 403.

[0040] In a procedure 403, it is considered that the following frame image $In+1$ of a current processing-object frame image is the new processing-object frame image In . And a control signal is sent to videodisk equipment 1 or VTR2, the processing-object frame image In after updating is reproduced, and the frame image In is captured from a frame memory 3.

[0041] In a procedure 404, the tail of the shot which is the part continuously photoed in time with one video camera is detected. When the tail of a shot is detected, after considering that this shot is the new processing-object shot SH_k (k is one or more integers), it progresses to a procedure 405. When the tail of a shot is not detected, it progresses to a procedure 405, without doing anything. In addition, after progressing to a procedure 405, it is that this procedure 404 is performed again, after

incorporating the following processing-object frame image In+1. Therefore, when the tail of a shot is undetectable, after incorporating degree processing-object frame image In+1, detection of the tail of a shot will be tried again.

[0042] In addition, you may consider that one scene of a scenario etc. is a shot. Therefore, all the "shots" used by the explanation after a procedure 404 may be reset on "one scene of a scenario."

[0043] Moreover, the image change model method introduced by the Prior art is proposed as the detection approach of the tail of a shot. Furthermore, in this example, although the image is automatically divided into the shot, while a user looks at an image etc., you may divide. A procedure 404 may be skipped when the image is beforehand divided into the shot.

[0044] After a procedure 405, even a procedure 408 is the image epitome approach using the following three assumptions built based on the subjectivity evaluation result of a rapid-traverse image, and reproduction speed is determined that it sees on "subjectivity target and reproduction speed will extinguish a quick part and late partial."

[0045] - In the part (it is hereafter called the long duration similar section) which the image which carried out long duration resemblance follows, it is sensed that reproduction speed is slow.

[0046] - In the part (it is hereafter called the high-speed operation section) of the image containing the body which moves at high speed, it is sensed that reproduction speed is quick.

[0047] - When an image is summarized for every contents and it divides into a shot, it is sensed at a shot (it is called a short-time length shot) with short display time that reproduction speed is quick.

[0048] In addition, the epitome image created by this approach will be called a section gear change epitome image. Moreover, in this example, the time amount sections other than the long duration similar section and the high-speed operation section will be called the standard section. Furthermore, the reproduction speed of the standard section is set as 4X.

[0049] A procedure 405 is processing which detects the long duration similar section. This detection processing is performed because the image of 4X rapid-traverse playback samples the image of four-frame spacing and investigates the similarity between that image, since it consists of images of four-frame spacing. The concrete actuation in a procedure 405 is described using drawing 5 .

[0050] In a procedure 501, the frame number of the processing-object frame image In is investigated first. Next, this frame image judges whether it is contained in the image sequences of four-frame spacing. Since the long duration similar section is detected only using the image of four-frame spacing, when judged with "It is contained in image sequences", in order to detect the long duration similar section, it progresses to a

procedure 502. When judged with "It is not contained in image sequences", a procedure 405 is ended and it progresses to the procedure 406 of drawing 4.

[0051] In a procedure 502, the similarity S between frame image In-4 of the image sequences in front of one (n, n-4) is calculated by the processing-object frame image In and the above-mentioned image sequences. However, the count approach of the similarity between frame images was decided so that correlation with the motion which the user who looks at an epitome image can ignore subjectively might be lost.

[0052] Here, the count approach of the similarity between frame images is described briefly. Although various kinds of approaches, such as chi2 assay (Nagasaka, Tanaka, "the method of detecting a scene change of a video work automatically", the 40th time national conference of Information Processing Society of Japan, 1Q-1990 [5 or]), can be considered In this example, a common color ratio method (a Yamada, Fujioka, Kanamori, Matsushima, and "examination of scene change detection approach of having observed common color for every subregion" television society technical report, September, 1993, Vol.17, No.55, pp 1-6), and the same approach are used.

[0053] With the static image and the image only including the motion which can be disregarded subjectively, the distance which the body on a screen moves to 4 frame time on a screen should become very small compared with the width of face of a screen. According to experience of writers, this distance becomes 2% or less of the width of face of a screen. Then, as shown in drawing 6, the frame image In is divided into Subregion R (j, n) (however, j or more 1 16 or less integer). At this time, the histogram of a color hardly changes between a static image, the image In only including the motion which can be disregarded subjectively, and the subregions R (j, n) and R (j, n-4) where In-4 correspond. In addition, although the number of subregions was set to 16 in this example, it is not necessary to be necessarily 16.

[0054] On the other hand, if change of a shot etc. occurs between the frame image In and In-4, between In and the subregion where In-4 correspond, the color which a photographic subject changes and constitutes subregion will change. Then, this similarity is calculated so that the similarity Sp (j, In, In-4) of subregion which corresponds according to the area of the color which appeared newly, and the color which disappeared between corresponding subregions may decrease.

[0055] And the average of the frame image In and the similarity of the subregion where In-4 correspond is calculated, and it is regarded as the similarity S of a frame image (n, n-4). Namely, [0056]

[Equation 1]

$$S (n , n - 4) = \sum_{j=1}^{16} S_p (j , I_n , I_{n-4}) \cdots \cdots \cdots (1)$$

It calculates.

[0057] A procedure 503 is processing which updates the head frame number of the long duration similar section. The image in the long duration similar section is [whether since it is mutually similar, the similarity S between the images of time series $(n - 4)$ becomes more than threshold θ_{astill} , and] conditional expression. $S(n - 4, n) \geq \theta_{\text{astill}}$ (2)

*****. (2) When a formula is materialized, consider that $(n - 4)$ is the middle of the long duration similar section, and if the similar section edge N_b showing a head frame number is not set up, substitute $(n - 4)$ for N_b .

[0058] A procedure 504 is processing which detects the candidate of the long duration similar section. It is [whether the similarity of the image of four frame spacing becomes more than the above-mentioned threshold θ_{astill} between the frame image of the similar section edge N_b , and the processing-object frame image I_n , and] conditional expression. $S(N_b + (i - 1) \times 4, N_b + i \times 4) \geq \theta_{\text{astill}}$, and $1 \leq i \leq (n - N_b) / 4$.. (3)

*****. (3) When a formula is materialized, consider that between the frame image I_{N_b} of a similar section edge and the above-mentioned processing-object frame images I_n is the candidate of the long duration similar section, and progress to a procedure 505. (3) Since the candidate of the similar section cannot be determined for a long time when a formula is not materialized, progress to the procedure 406 of drawing 4 .

[0059] "The part which carries out loose image change" as shown in drawing 7 other than "the part which a similar image follows" is contained in the candidate who asked in the procedure 504. It is [whether in a procedure 505, the similarity between a long duration similar section candidate's head frame image I_{N_b} and other frame images becomes first more than threshold θ_{aratio} , and] conditional expression. $S(N_b, N_b + i \times 4) \geq \theta_{\text{aratio}}$, $1 \leq i \leq (n - N_b) / 4$ (4)

*****. Next, it is [whether when (4) types are materialized, this candidate's time amount length is long compared with the minimum time amount length T_{still} of the long duration similar section, and] conditional expression. $n - N_b \geq T_{\text{still}}$ (5)

*****. However, the minimum time amount length T_{still} in (5) types sets up at the time of activation of the procedure 401 of drawing 4 .

[0060] When (4) and (5) types are materialized in coincidence, after considering that the candidate who asked in the procedure 504 is the long duration similar section, the procedure 405 described above is ended.

[0061] (4) If the image after the processing-object frame image I_n is not processed when a formula is materialized and (5) types are not materialized, the candidate who asked in the procedure 504 cannot judge whether it is a part of long duration similar section. Then, a procedure 405 is ended, without doing anything.

[0062] (4) When a formula is not materialized, consider that this candidate is "the part

which carries out loose image change.” Moreover, since (2) types are materialized, it is considered that a frame number (n-4) or subsequent ones is the candidate of the long duration similar section. (n-4) is substituted for the similar section edge Nb showing the head frame number of the candidate of the long duration similar section. [0063] In addition, in this example, the count approach of the similarity between frame images was decided so that correlation with the motion which the user who looks at an epitome image can ignore subjectively might be lost. However, since a user’s subjectivity varies by the user, it may think as important a motion of the body which a user observes and may determine the count approach of similarity.

[0064] Moreover, in this example, it asked for “the part which a similar image follows” using (3) and (4) types. That is, the similarity between the images of time series and the similarity on the basis of the head frame image of the similar section were used. However, since the same effectiveness is acquired even if it uses the approach on the basis of the frame image of the last of the similar section etc., “the part which a similar image follows” may be calculated using approaches other than (3) and (4) types. Moreover, “the part which a similar image follows” may be calculated only using (4) types.

[0065] It progresses to a procedure 406 after procedure 405 termination of drawing 4 . A procedure 406 is processing which detects “the part of the image containing the body which moves at high speed.” The concrete actuation in a procedure 406 is described using drawing 8 .

[0066] A procedure 801 is processing which detects the pixel on the processing-object frame image In and the body which moves at high speed using image In-1 in front of that. The principle of this processing is explained using drawing 9 .

[0067] As shown in drawing 9 , when the body on a location p (however, p vector) carries out high-speed migration, within distance Lmin, it is assumed between frame image In-1 and In that the body of a basis does not already exist. At this time, by migration of a body, on In-1 The brightness f of the pixel of a location p (p, n-1) Brightness f of the pixel in the circle of the radius Lmin centering on the location p same on In (p+d) All the brightness differences between n) (however, vector in which d has the die length below Lmin) become thetaw1 or more thresholds set up beforehand.

$$|f(p, n-1) - f(p+d, n)| \geq \text{thetaw1}, |d-d| \leq L_{\min} \times L_{\min} \dots (6)$$

It *****. However, in (6) types, |a| expresses the absolute value of scalar quantity a, and (d-d) expresses the inner product of Vector d.

[0068] Furthermore, a noise is removed by detecting only a larger body than the circle which makes Rmin a radius. As it considers that the same brightness field is the same body and was shown in drawing 9 , it is the brightness f of the pixel of a location p (p+D) on In-1. Conditional expression from which the brightness difference between n-1 (however, vector in which D has the die length below Rmin) becomes less than

thetaw1 threshold $|f(p, n-1) - f(p+D, n-1)| < \text{thetaw1}$ and $|D-D| \leq R_{\text{min}} \times R_{\text{min}}$.. (7)

When *****ing), it investigates whether (6) types are materialized in a location p.

[0069] After detecting the pixel p which fills (6) and (7) types with the procedure 801 of drawing 8 to coincidence based on the above thing, it considers that this pixel is "a pixel on the body which moves at high speed", and it is registered as a high-speed operation pixel.

[0070] in addition, the case where pixel infanticide is carried out in order to simplify processing and 8pixelx -- it may be considered that the average of the pixel for eight lines is one pixel Thus, in making one pixel represent the contents of two or more pixels, the area per pixel becomes large like a mosaic. In this case, as shown in drawing 10 , you may consider that the pixel which fills only (6) types is a high-speed operation pixel.

[0071] Moreover, the approach using the motion vector as an approach of detecting the body which moves at high speed can be considered. Like the following procedure, a high-speed operation pixel may be detected using a motion vector.

[0072] After dividing into two or more blocks [example of procedure of detecting body which moves at high speed using motion vector] procedure 1 image, it asks for the motion vector for every block using the image of two sheets of time series.

[0073] The block with which the magnitude of procedure 2 motion vector becomes more than constant value is searched for.

It is considered that the pixel contained in the block searched for in procedure 3 procedure 2 is a high-speed operation pixel.

[0074] Furthermore, although the pixel in comparison with Pixel p was made into the pixel contained in a specific circle by (6) and (7) formulas, as shown in drawing 11 , it is good also as a pixel contained in the square containing Pixel p.

[0075] Moreover, a high-speed migration pixel may be detected out of the pixel contained to the field assumed beforehand supposing the field which a user observes. For example, as shown in drawing 12 , it may assume that a user always observes near the center of a screen, and a high-speed migration pixel may be detected.

[0076] In a procedure 802, the total of the high-speed operation pixel contained in frame image In-1 is investigated. And when the total of a high-speed operation pixel becomes a large value compared with threshold thetam set up in order to remove a noise, it is considered that between this frame image In-1 and following frame image In is a high-speed operation section candidate.

[0077] If it moves by the result of subjectivity evaluation of a rapid-traverse image only for a moment even if the body which moves at high speed exists, it is ignored subjectively. Then, a procedure 804 excepts the section which became a candidate by momentary motion out of a high-speed operation section candidate. Although some implementation approaches of this processing are considered, they describe that one approach here.

[0078] In this example, since the image of 4X rapid-traverse playback has been criteria, an image is summarized every four frames and it will be called a section unit. In a procedure 803, it judges whether the high-speed operation section candidate was detected by the procedure 802. When it judges with "It was detected", it supposes that it writes UN [a section unit including this high-speed operation section candidate] (N is one or more integers), and progresses to a procedure 804. Since a high-speed operation section candidate does not exist when it judges with "It is not detected", a procedure 406 is ended.

[0079] A procedure 804 is processing which judges whether it continues more than the lower limit Nmove that the section unit including a high-speed operation section candidate set up beforehand. Moreover, it investigates whether all the section units between the section unit UN and section unit $UN - (Nmove - 1) * 4$ in front of the $(Nmove - 1)$ individual contain a high-speed operation section candidate. Next, when judged with "A high-speed operation section candidate is included", it is considered that between frame image $IN - Nmove * 4$ and $IN(s)$ is the high-speed operation section.

[0080] The procedure 406 stated above is ended after termination of a procedure 804. And it progresses to a procedure 407 after termination of the procedure 406 of drawing 4.

[0081] A procedure 407 is processing which determines the image reproduction rate for a section gear change epitome image. First, the reproduction speed of the high-speed operation section is determined as 2X. Next, what does not include the high-speed operation section in the middle is taken out out of the long duration similar section, and the reproduction speed is determined as 8X. Furthermore, it considers that the remaining section which reproduction speed has not determined is the standard section, and the reproduction speed is determined as 4X.

[0082] In addition, in a procedure 407, although the reproduction speed of the high-speed operation section, the standard section, and the long duration similar section was set as 2X, 4X, and 8X, respectively, as long as the reproduction speed of the high-speed operation section is slow compared with the standard section and the reproduction speed of the long duration similar section is quick compared with the standard section, you may set it as other values. Moreover, reproduction speed may be determined using display time like the case which reproduces the whole long duration similar section in 1 second.

[0083] When an image is reproduced according to the reproduction speed determined in the procedure 407, display time becomes under the shot length lower limit Nshot set up beforehand at some shots. Then, reproduction speed is corrected in a procedure 408.

[0084] In a procedure 408, the time amount NSHk required for playback (k is one or more integers) is first found from each shot SHk for which it asked in the procedure 404. Next, it is [whether it becomes under the shot length lower limit Nshot and]

conditional expression. $NSHk < Nshot$ (8)

*****. (8) When a formula is materialized, consider that Shot SHk is a short-time length shot, and determine reproduction speed so that the display time of a shot may become the shot length lower limit Nshot.

[0085] In addition, in this example, although the shot length lower limit was made into the fixed value, a shot length lower limit may be changed according to the contents, such as fineness of a texture.

[0086] It progresses to a procedure 409 after termination of a procedure 408. A procedure 409 is a procedure for realizing the image epitome approach which reproduces the shot of time series one after another, reminding this "rhythm", after determining reproduction speed that the boundary of a shot and the rhythm of the period set up beforehand will have correlation. Reproduction speed is determined so that first an epitome image may have the rhythm of the period set up beforehand on the boundary of a shot. Next, holding a rhythm, reproduction speed is corrected so that a user can grasp the contents of each shot. In addition, the epitome image created in the procedure 409 will be called a rhythm presentation epitome image.

[0087] The concrete actuation in a procedure 409 is described using drawing 13. What is necessary is just to determine reproduction speed so that display time of for example, each shot may be made regularity in order to give the rhythm of the period beforehand set as the boundary of a shot. Moreover, as an approach of reminding a rhythm, various kinds of approaches, such as an approach of accustoming a sound on the boundary of a shot, and the approach of displaying a static image on the boundary of a shot, can be considered.

[0088] At this example, like the example shown in drawing 14, after displaying the head frame image of a shot between the quiescence time amount Tstl, the whole shot is indicated by rapid traverse by the rapid-traverse time amount Tlen. The rhythm of spacing (Tstl+Tlen) occurs at this time. For example, if Tstl is made into 0.2 seconds and Tlen is set as 0.8 seconds, the rhythm of 1-second spacing will occur.

[0089] In the procedure 1301 of drawing 13, frame number LSHk (k is one or more integers) of the processing-object shot SHk is calculated.

[0090] in a procedure 1302, the display time of the rapid-traverse part of the processing-object shot SHk turns into the rapid-traverse time amount Tlen set up beforehand -- as -- $V(k) = LSHk / Tlen$ (9)

Reproduction speed V (k) is decided to be alike.

[0091] In a procedure 1303, the rate upper limit Vmax (k) which is the criteria of the reproduction speed decision of the processing-object shot SHk is determined so that a user can grasp the contents of the processing-object shot SHk. The decision approaches of this rate upper limit Vmax (k) are various idea ****. For example, the reproduction speed determined in the procedure 407 is averaged in the processing-object shot SHk, and how to assign this value to the rate upper limit Vmax

(k) can be considered. Moreover, the method of considering that the fixed value set up beforehand is $V_{\max}(k)$ is also considered. Furthermore, the following method of having used the similar section and standard section and the high-speed operation section for a long time which was determined in the procedure 407 is also considered.

[0092] At the shot which does not include the high-speed operation section, the rate upper limit $V_{\max}(k)$ is set as a large value compared with the shot of only the standard section, including the [rate upper-limit decision approach using long duration similar section, standard section, and the high-speed operation section] procedure 1 long-duration similar section. For example, it is set as 8X.

[0093] At the shot which does not include the long duration similar section, the rate upper limit $V_{\max}(k)$ is set as a small value compared with the shot of only the standard section, including the procedure 2 high-speed-operation section. For example, it is set as 2X.

[0094] At a shot including both the procedure 3 high-speed-operation section and the long duration similar section, and the shot of only the standard section, the rate upper limit $V_{\max}(k)$ is set as the same value. For example, it is set as 4X.

[0095] It is [whether in a procedure 1304, reproduction speed $V(k)$ determined by (9) formulas becomes more than the rate upper limit $V_{\max}(k)$ determined in the procedure 1303, and] conditional expression. $V(k) \geq V_{\max}(k)$ (10)

*****. (10) When a formula is materialized, in order to make a user grasp the contents, extend the display time of the rapid-traverse part of the processing-object shot SHk, and correct reproduction speed. This correction processing is performed in a procedure 1305.

[0096] (10) When a formula is not materialized, end the procedure 1309 stated above like the part of the shot SH1 of drawing 15 after determining the reproduction speed of the rapid-traverse part of Shot SHk as $V(k)$.

[0097] As already stated, a procedure 1305 is processing for making a user grasp the contents of the shot, extends the display time of a shot and corrects reproduction speed.

[0098] in order to give a rhythm to a shot boundary -- the display time of a shot -- “-- others -- it is made integral multiple” of a shot. However, in order to reproduce each shot one after another for a short time, display time of the rapid-traverse part of each shot is shortened as much as possible. Then, the display time of the rapid-traverse part of the shot which is satisfied with a procedure 1305 of (10) types is set as $2 \times T_{\text{len}}$ which is twice other shots.

[0099] Next, the decision approach of reproduction speed is described. (10) Correct so that the reproduction speed for the first portion of the shot which fills a formula may become below the rate upper limit $V_{\max}(k)$ determined in the procedure 1303. Namely, like the shot SH2 of drawing 15, after setting the reproduction speed V_P for the first portion (k) as the value below the rate upper limit $V_{\max}(k)$, it is determined in

the second half that the reproduction speed VR of a part (k) will set display time of the whole rapid-traverse part to 2xTlen. However, the reproduction speed VR of a part (k) may exceed Vmax (k) in the second half. Moreover, in order to hold a rhythm, a stationary part is put in between parts in a part for the first portion, and the second half of a rapid-traverse part.

[0100] The reproduction speed VP in the first half of a rapid-traverse part (k) is [0101].

[Equation 2]

$$V P (K) \left\{ \begin{array}{l} V_{\max} (K) , \frac{V (K)}{2} > V_{\max} (K) \text{ のとき} \\ \frac{V (K)}{2} , \frac{V (K)}{2} \leq V_{\max} (K) \text{ のとき} \end{array} \right. \dots\dots\dots (11)$$

It can determine "To be alike." Moreover, the reproduction speed VR in the second half of a rapid-traverse part (k) is [0102].

[Equation 3]

$$V R (K) = \frac{L S H_k - V P (K) \times T l e n}{T l e n} \dots\dots\dots (12)$$

It can determine "To be alike." The procedure 409 stated above is ended after count of (11) and (12) types.

[0103] In addition, when some shots which cannot grasp the contents may occur, a procedure 409 may be ended after termination of a procedure 1302. Moreover, when change of a shot is not predicted, reproduction speed V (k) may be determined, without using (9) types.

[0104] It progresses to a procedure 410 after termination of the procedure 409 of drawing 4 . After from the procedure 410 to the procedure 412 unifies a similar shot, regards it as a shot group using correlation between the shots of time series and chooses from one shot group the partial dynamic image which consists of a frame image (for example, 1 second) of two or more time series, it is a procedure for realizing the image epitome approach which reproduces the partial dynamic image of time series one after another. First, correlation between the shots of time series is used, a similar shot is unified, and a shot group is created. Next, every one partial dynamic image is chosen from each shot group, and it is processed so that the description of a shot group may be specified. For example, in unifying a shot including a prototype object and creating a shot group, this prototype object serves as the description of a shot group. Moreover, when time amount length unifies an almost equal shot, this time amount length serves as the description of a shot group.

[0105] When reproducing an epitome image, the processed partial dynamic image is reproduced one after another. In addition, the selected partial dynamic image may be

reproduced directly, without performing processing.

[0106] The epitome image created with the procedure 412 from the procedure 410 will be called a partial selection epitome image.

[0107] A procedure 410 is processing which creates a shot group. The concrete actuation in a procedure 410 is described using drawing 16 .

[0108] Since a shot is the unit of the contents, the frame image in one shot has one common description. For example, a person appears in all frame images at the shot which follows a person. Therefore, the contents of the shot can be represented using some frame images in a shot. In order to simplify processing, the number of sheets Nrep of a frame image which represents the contents is set up beforehand, and the contents of the shot are represented with this example using a Nrep sheet [in a shot] frame image (it is hereafter called the image between representation space-time).

[0109] In a procedure 1601, the image representing Shot SHk is determined, the images Ik and j between representation space-time and (k regard it as one or more integers, and j regards it as integer) below or more 1Nrep. However, as the decision approach of the image between representation space-time, they are various idea ****, such as the approach of choosing the Nrep sheet of a head part, and the approach of choosing the image of a fixed time interval. Moreover, as long as there is no limit in throughput, you may consider that all the frame images in a shot are the images between representation space-time.

[0110] The shot of the similar contents of time series is unified after termination of a procedure 1601 using the image between representation space-time, and it is regarded as a shot group. As an example of the shot of similar contents, there are a shot of a similar background, a shot of a similar photographic subject, etc. In this example, a shot is unified paying attention to the color of the body on a screen, and a motion.

[0111] According to analysis of writers, by the image between representation space-time of the shot of similar contents, the body in which the color common to each image between representation space-time exists, and has a common color carries out a common motion. Although it is difficult to investigate whether two bodies which appear on a screen are carrying out the common motion, the change by motion of a body is easily detectable using a pixel change field, a motion vector, etc. which are mentioned later. Then, the image between representation space-time is processed and it divides into the field (it is hereafter called a motion area) influenced of change by motion of the body on a screen, and the remaining fields (it is hereafter called a quiescence field). At this time, the shot of similar contents satisfies the following common color ratio conditions.

[0112] Common color ratio conditions: When calling the color which constitutes the quiescence field of the image between representation space-time, and the color which constitutes a motion area a quiescence configuration color and *****,

respectively, the quiescence configuration color and ***** which it is under [between representation space-time / image] community exist, and a pixel with these colors accounts for the rate more than threshold θ_{shot} by each image between representation space-time.

[0113] In addition, the example of the shot of similar contents which satisfies common color ratio conditions is shown below.

Example 1: a shot of a similar background like [in case the shot photoed in the same location continues]. two shot SH_{k-1} of the scenery of the poolside shown in drawing 17 , and image I_k[of SH_k] between representation space-time- in 1, 1, I_k, and 1 The quiescence configuration color D in rate =15%I_{k-1} of the quiescence configuration color C in rate =20%I_{k-1} of the quiescence configuration color B in rate =60%I_{k-1} of the quiescence configuration color A in I_{k-1} and 1 and 1 and 1 and 1 comparatively =5%I_k, The quiescence configuration color C in I_k and 1 comparatively 80% The rate of the quiescence configuration color A in one = =15%I_k, I_k of the quiescence configuration colors A and C in which the rate in I_{k-1} [of the quiescence configuration color E in one] of the quiescence configuration colors A and C which become =5% comparatively and are common, and 1 is common sum =75%, the sum of the rate in one = since it becomes 95%, a common quiescence configuration color occupies 75% or more of the whole screen.

[0114] Example 2: a shot of a similar photographic subject like [in case the shot of the same photographic subject continues]. two shot SH_{k-1} of tailing of the automobile shown in drawing 18 , and image I_k[of SH_k] between representation space-time- in 1, 1, I_k, and 1 The quiescence configuration color Y in rate =10%I_{k-1} of the quiescence configuration color X in rate =40%I_{k-1} of the quiescence configuration color W in rate =40%I_{k-1} of the quiescence configuration color V in I_{k-1} and 1 and 1 and 1 and 1 comparatively =10%I_k, The quiescence configuration color W in I_k and 1 comparatively 40% The rate of the quiescence configuration color Z in one = =35%I_k, Rate =10%I_k of the quiescence configuration color X in one, the rate of the quiescence configuration color Y in one = It becomes 15%. I_k of ***** W, X, and Y in which the rate in I_{k-1} of common ***** W, X, and Y and 1 is common sum =60%, the sum of the rate in one = since it becomes 60%, common ***** occupies 60% or more of the whole screen.

[0115] In the procedure 1602 of drawing 16 , a motion area is detected from the image between representation space-time, and the image between representation space-time is divided into a motion area and a quiescence field. When a motion vector is used, a motion area comes to be shown in drawing 19 . Moreover, since the brightness of the pixel in the same location is almost equal when a body does not move between two frame images, when it considers that the next pixel change field is a motion area, a motion area comes to be shown in drawing 20 .

[0116] pixel change field: — the frame images I_k and j in the image between representation space-time, and deflection after the 1 frame time between — MU

images -- " -- the absolute value of the brightness difference of the pixel in the same location -- threshold [] -- set of pixel" which becomes thetaW1 or more.

[0117] A procedure 1603 is processing which asks for a quiescence configuration color and *****. The red component of a pixel, a green component, and a blue component will be hereafter called R, G, and B, respectively, and a procedure is explained. First, it asks for the histogram HS (c, SHk) of the 512 colors c of RGB eight gradation each in the quiescence field of the image between representation space-time of Shot SHk. Similarly, it asks for the histogram HM (c, SHk) of a motion area. Next, in order that the number of pixels may consider that few colors are noises and may except them, number of minimum pixels thetaH of a configuration color is set up, and it asks for the quiescence configuration color HSVk and ***** HMV_k (k is one or more integers) using a degree type.

[0118]

HSV_k = {c | HS(c, SH_k) > thetaH} (13)

HMV_k = {c | HM(c, SH_k) > thetaH} (14)

Below, the shot of the head of a shot group is described to be SH_{top} (for top to be one or more integers). A procedure 1604 to the procedure 1606 is the processing which investigates the shot of the time series after this shot SH_{top}, and investigates the shot to where fulfills common color ratio conditions.

[0119] In a procedure 1604, it supposes that the quiescence configuration color which is common between shot SH_{top}–SH_{top}+m, and ***** are called the quiescence common colors HSC_{top} and m and the ** common colors HMC_{top} and m (top and m are one or more integers), respectively, and they are calculated using a degree type.

[0120]

[Equation 4]

$$HSC_{top,m} = \{c \mid c \in HSV_{top} \cap HSV_{top+1} \cap \dots \cap HSV_{top+m}\} \dots (15)$$

$$HMC_{top,m} = \{c \mid c \in HMV_{top} \cap HMV_{top+1} \cap \dots \cap HMV_{top+m}\} \dots (16)$$

In a procedure 1605, it supposes that "the rate that a pixel with the quiescence common colors HSC_{top} and m and the ** common colors HMC_{top} and m occupies in the image between representation space-time of Shot SH_k" is called the image common color ratio AMC (k, top, m), and calculates using a degree type.

[0121]

[Equation 5]

$$AMC(k, top, m) = \frac{\sum_{c \in HSC_{top, m}} HS(c, SH_k) + \sum_{c \in HMC_{top, m}} HM(c, SH_k)}{\sum_{c \in HSV_k} HS(c, SH_k) + \sum_{c \in MV_k} HM(c, SH_k)} \dots\dots (17)$$

Formula which expresses formation of common color ratio conditions with a procedure 1606 $AMC(k, top, m) \geq \theta_{shot, top} \leq k \leq top+m \dots\dots\dots (18)$

When materialized to ** and all k, it is considered that shot $SH_{top} - SH_{top+m}$ satisfies common color ratio conditions. In order to investigate the shot to where fulfills common color ratio conditions, when (18) types are materialized to all k, after adding 1 to m, it returns to a procedure 1604. When that is not right, it progresses to a procedure 1607.

[0122] In this example, it asks for the tail of a shot group using the image common color ratio between two shots of time series. First, the image common color ratios $AMC(k, k1)$ and $AMC(k+1k, 1)$ (k is an integer below $top+m$ more than top) which (17) types defined are calculated in a procedure 1607. Former $AMC(k, k1)$ expresses "the rate that two shots SH_k , and the quiescence common color and ** common color of SH_{k+1} occupy in the image between representation space-time of the front shot SH_k ." Latter $AMC(k+1k, 1)$ expresses "the rate that two shots SH_k , and the quiescence common color and ** common color of SH_{k+1} occupy in the image between representation space-time of next shot SH_{k+1} ."

[0123] By the procedure 1608, it is the minimum value of an image common color ratio. $SS(k) = \min(AMC(k, k1), AMC(k+1k, 1)) \dots\dots\dots (19)$

It calculates and is regarded as the priority for integration of Shot SH_k and SH_{k+1} .

[0124] In a procedure 1609, it asks for the tail of a shot group out of shot $SH_{top} - SH_{top+m-1}$ which fulfills common color ratio conditions, and its following shot SH_{top+m} . Formula which expresses that a priority is min out of the priorities $SS(top) - SS(top+m-1)$ of integration between these shots $SS(k_{min}) \leq SS(k)$, $top \leq k \leq top+m-1 \dots\dots\dots (20)$

***** $SS(k_{min})$ is calculated and it is considered like drawing 21 (a) and the example of (b) that the shot $SH_{k_{min}}$ from which a priority becomes min is the tail of a shot group. ((b) drawing 21 (a) and 2nd shot) And the procedure 1610 stated above is ended.

[0125] In addition, the approaches of asking for a shot group are various idea ****. For example, at the shot of time series, since the contents are similar when time amount length is almost equal, it can be regarded as one shot group. Therefore, you may ask for a shot group using a different approach from this example.

[0126] Moreover, it may consider after a procedure 1606 that tail SH_{k+m-1} of the shot which fulfills common color ratio conditions is the tail of a shot group, without

progressing to a procedure 1607, and a procedure 410 may be ended.

[0127] It progresses to a procedure 411 after termination of the procedure 410 of drawing 4 . Every one partial dynamic image is chosen from each shot group in a procedure 411. However, the selection approach of a partial dynamic image has various kinds of approaches, such as a head part of the head shot of a shot group, and a head part of the shot representing the contents of the shot group. Moreover, various kinds of approaches, such as the approach of making it into a fixed length and a method of changing time amount length according to the audio contents, can be considered also about the setting approach of the time amount length of a partial dynamic image.

[0128] In addition, in a procedure 411, although the partial dynamic image was chosen, you may choose one static image from each shot group at a time. In this case, it will become an epitome image if the selected static image is displayed one after another. Moreover, the selected static image may be reduced and you may indicate by list. A list display result can be treated on a par with an epitome image.

[0129] It progresses to a procedure 412 after termination of a procedure 411. A procedure 412 is processing for specifying the description of a shot group, and carries out processing processing of the partial dynamic image chosen in the procedure 411.

[0130] The shot contained in a shot group fills common color ratio conditions with this example. That is, the quiescence common color and ** common color for which each shot which constitutes a shot group was asked in the procedure 1607 exist.

[0131] Since a quiescence common color and a ** common color express the color of the body which exists in common with a shot group, it can be considered that they are the description of a shot group. For example, if it is the shot of a similar background, the color of a background is expressed, and if it is the shot of a similar photographic subject, the color of a photographic subject is expressed. So, in a procedure 412, the pixel of a quiescence common color and a ** common color is similarly displayed as usual, and pixels other than this are expressed as half brightness. In addition, as long as it emphasizes a quiescence common color and a ** common color, what kind of processing art may be used, such as making the brightness of the pixel except a quiescence common color and a ** common color into zero.

[0132] Moreover, the processing art of an image is decided in accordance with the integrated approach of a shot, and various variations exist. For example, when time amount length unifies an almost equal shot, the processing art which displays the mean-time length of a shot on the partial dynamic-image bottom can be considered like drawing 22 .

[0133] It returns to a procedure 402 after termination of a procedure 412. In addition, when the description of a shot group does not need to be specified, you may return to a procedure 402 after termination of a procedure 411, without performing a procedure 412.

[0134] The above is one example of image epitome processing of the 1st computer 5 in drawing 1 , and is detailed explanation of the procedure 201 of drawing 2 .

[0135] The procedure 202 of drawing 2 is performed after termination of image epitome processing. The 1st calculating machine 5 in drawing 1 controls videodisk equipment 1 and VTR2 by the procedure 202, and reproduces in it a processing-object image and the image summarized in the procedure 201. The reproduced image is compressed by the MPEG method in the image compression equipment 4 of drawing 1 , and is memorized by the file server 6.

[0136] However, since the procedure 201 of this example creates the playback approach of an epitome image as shown in drawing 3 , without creating the summarized image, it is controlling videodisk equipment 1 and VTR2 according to the playback approach of this epitome image, and performs playback of an epitome image.

[0137] In addition, it is not necessary to necessarily compress by the MPEG method, and you may compress by other compression methods, such as a JPEG method. Moreover, not only the approach of compressing an image directly about the store method of an epitome image, as already stated but other approaches are various idea
****.

[0138] Below, epitome image regeneration of the 2nd computer 7 of drawing 1 which is concrete actuation of the procedure 203 in drawing 2 is described.

[0139] Drawing 23 is the flow chart of one example of epitome image regeneration of the 2nd calculating machine 7 in drawing 1 .

[0140] The graphic display approach is chosen in a procedure 2301. However, much alternative can be considered. For example, what is necessary is just to choose "standard speed playback of a processing-object image" to see the contents of the image in detail. The image which a user wants to see at the time of graphic display is reproduced at an ordinary rate.

[0141] Moreover, what is necessary is just to choose the "section gear change epitome image" created by processing from the procedure 405 of drawing 4 to a procedure 408 to grasp the contents focusing on actuation of a photographic subject. The whole image is reproduced seeing subjectively and making it reproduction speed enter in tolerance at the time of graphic display.

[0142] In addition, there is individual difference in subjectivity. Therefore, a user may choose the reproduction speed decision approach according to his subjectivity. Instead of performing the following three processings, only the part may be performed or, specifically, the processing "which is the part of the fine image of a texture and makes reproduction speed slow relatively" may be added.

[0143] - the procedure 408 for making reproduction speed late relatively at procedure 406 and the short-time merit shot for making reproduction speed late relatively in procedure 405 and the high-speed operation section for making reproduction speed quick relatively in the similar section for a long time -- what is necessary is just to

choose the "rhythm presentation epitome image" created with the procedure 409 of drawing 4 to know further the time flow of the contents included in an image. The shot which summarized the part of the same contents at the time of graphic display is reproduced one after another.

[0144] What is necessary is just to choose the "partial selection epitome image" created by processing from the procedure 410 of drawing 4 to a procedure 412 to see the part of different contents as much as possible little by little. The shot group which summarized the shot of the contents from which time series is similar at the time of graphic display is reproduced one after another.

[0145] In the procedure 2302 of drawing 23, according to the selection result by the procedure 2301, a required image is called from the file server 6 of drawing 1, and the image is reproduced. However, when only the frame number information for epitome image creation is saved at the file server 6, it reproduces according to frame number information, summarizing a processing-object image. Moreover, a user may interrupt playback at the favorite time, specifying "to [from which frame] which frame it reproduces", specifying "from which frame it reproduces", and looking at an image before image reproduction.

[0146] In addition, what is necessary is just to decide the flow of the image epitome processing shown in drawing 4 in accordance with the alternative of a procedure 2301. For example, when a "rhythm presentation epitome image" is not included in the alternative of a procedure 2301, it is not necessary to perform the procedure 409 of drawing 4. Similarly, when a "section gear change epitome image" is not included, it is not necessary to perform from the procedure 405 of drawing 4 to the procedure 408, and when a "partial selection epitome image" is not included, it is not necessary to perform from the procedure 410 of drawing 4 to the procedure 412.

[0147] Moreover, although this example showed the example which reproduces the summarized image and a processing-object image on the 2nd calculating machine 7 of drawing 1, you may reproduce on the alien machine connected to the file server 6. the case of drawing 1 where may reproduce on the 1st calculating machine 5 and two or more calculating machines are connected to the file server 6 -- those the count of all -- it is on board and you may reproduce.

[0148] As mentioned above, according to the section gear change epitome image of this example, the time amount section when an image similar over beyond fixed time amount continues is detected, and it is regarded as the long duration similar section, and by reproducing the long duration similar section quickly relatively, since an image changes one after another for a short time, section <TXF FR=0001 HE=250 WI=080 LX=0200 LY=0300> part" with slow reproduction speed disappears on "subjectivity target. Moreover, the time amount section containing the body which moves at high speed is detected, and it is regarded as the high-speed operation section, and since the body which runs by making late relatively reproduction speed of the high-speed

operation section at high speed stops existing, "it is a part with quick reproduction speed subjectively" decreases. Furthermore, since each contents are displayed beyond fixed time amount by preparing a minimum in the display time length of each shot after summarizing an image for every contents and dividing into two or more shots, "it is a part with quick reproduction speed subjectively" disappears. Thus, compared with a conventional method, fatigue of a user is mitigated by reproducing the whole image, seeing subjectively and making it reproduction speed enter in tolerance. Moreover, contents centering on actuation of a photographic subject can be grasped.

[0149] Moreover, with the rhythm presentation epitome image of this example, after setting an upper limit as the display time length of each shot, the shot of time series is reproduced one after another, determining that reproduction speed will satisfy at least one side of the following two conditions, and reminding a rhythm.

[0150] Rhythm conditions: The boundary of a shot and the rhythm of the period set up beforehand have correlation.

[0151] Contents conditions : When the rate upper limit is set up so that it may guarantee "the part reproduced at the rate of below the rate upper limit which is the criteria of reproduction speed decision can surely grasp the contents", a part of [at least] reproduction speed of each shot becomes below a rate upper limit.

[0152] In summarizing an image as satisfying rhythm conditions, since change of a shot can be predicted, it can see, without overlooking no shots.

[0153] In summarizing an image as satisfying contents conditions, a user can grasp some of the contents at least in all shots. Since it is the shot which summarized the part of the same contents, the contents of all the shots can be grasped.

[0154] However, since change of a shot cannot be predicted when not satisfying rhythm conditions, an oversight of a shot occurs. When not satisfying contents conditions, the contents of some shots cannot be grasped. It is desirable that it is satisfied with coincidence of rhythm conditions and contents conditions.

[0155] Furthermore, with the partial selection epitome image of this example, after unifying a similar shot and creating a shot group using correlation between the shots of time series, the partial dynamic image chosen from each shot group is reproduced one after another. By using this epitome image, the part of different contents as much as possible can be seen efficiently.

[0156]

[Effect of the Invention] As mentioned above, this invention detects the time amount section when an image similar to the 1st over beyond fixed time amount continues, considers for a long time that it is the similar section, by reproducing the similar section quickly relatively for a long time, can change an image one after another in a short time, and can extinguish "it is a part with slow reproduction speed subjectively." Moreover, the time amount section of the image containing the body which moves at

high speed is detected, and it is regarded as the high-speed operation section, and by making late relatively reproduction speed of the high-speed operation section, the body which moves at high speed can be extinguished and it can decrease "it is a part with quick reproduction speed subjectively." Furthermore, after summarizing the part of the same contents to a shot, by preparing a minimum in the display time length of each shot, each contents can be displayed beyond fixed time amount, and it can extinguish "it is a part with quick reproduction speed subjectively." Thus, compared with a conventional method, fatigue of a user is mitigated by reproducing the whole image, seeing subjectively and making it reproduction speed enter in tolerance. Moreover, contents centering on actuation of a photographic subject can be grasped. [0157] Moreover, the shot of time series is reproduced one after another, determining that reproduction speed will satisfy at least one side of the following two conditions, and making the 2nd recollect a rhythm, after setting an upper limit as the display time length of each shot.

[0158] Rhythm conditions: The boundary of a shot and the rhythm of the period set up beforehand have correlation.

[0159] Contents conditions : When the rate upper limit is set up so that it may guarantee "the part reproduced at the rate of below the rate upper limit which is the criteria of reproduction speed decision can surely grasp the contents", a part of [at least] reproduction speed of each shot becomes below a rate upper limit.

[0160] In summarizing an image as satisfying rhythm conditions, since change of a shot can be predicted, it can see, without overlooking no shots.

[0161] In summarizing an image as satisfying contents conditions, a user can grasp some of the contents at least in all shots. Since it is the shot which summarized the part of the same contents, the contents of all the shots can be grasped.

[0162] However, since change of a shot cannot be predicted when not satisfying rhythm conditions, an oversight of a shot occurs. When not satisfying contents conditions, the contents of some shots cannot be grasped. It is desirable that it is satisfied with coincidence of rhythm conditions and contents conditions.

[0163] After unifying a similar shot and creating [3rd] a shot group using correlation between the shots of time series furthermore, the partial dynamic image chosen from each shot group is reproduced one after another. By using this epitome image, the part of different contents as much as possible can be seen efficiently.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] The whole image epitome equipment system chart in one example of this

invention

[Drawing 2] The flow chart of actuation of the image epitome equipment in this example

[Drawing 3] The conceptual diagram of an expression of the epitome image by the file format in this example

[Drawing 4] The flow chart of the image epitome processing in this example

[Drawing 5] The flow chart of the long duration similar section detection processing in this example

[Drawing 6] Drawing showing the example of creation of the subregion in this example

[Drawing 7] Drawing showing the part which carries out loose image change of the image in this example

[Drawing 8] The flow chart of the high-speed operation section detection processing in this example

[Drawing 9] Drawing showing the case where the body on the screen in this example moves to a high speed

[Drawing 10] Drawing showing the comparison of a pixel only using (6) types in this example

[Drawing 11] Drawing showing the comparison of the pixel of a location p in (6) in this example, and (7) types

[Drawing 12] Drawing showing the field which performs high-speed operation pixel detection in this example

[Drawing 13] The flow chart of the rhythm presentation epitome processing in this example

[Drawing 14] The conceptual diagram of the playback approach which fulfills the rhythm conditions in this example

[Drawing 15] The conceptual diagram of the playback approach of the rhythm presentation epitome image in this example

[Drawing 16] The flow chart of the shot integrated processing in this example

[Drawing 17] The conceptual diagram of the shot of the similar background in this example

[Drawing 18] The conceptual diagram of the shot of the similar photographic subject in this example

[Drawing 19] The conceptual diagram of the motion area in this example

[Drawing 20] The conceptual diagram of the motion area in this example

[Drawing 21] The conceptual diagram of the tail decision processing in this example

[Drawing 22] The conceptual diagram of the processing approach of the partial dynamic image in this example

[Drawing 23] The flow chart of the epitome image reproduction processing in this example

[Description of Notations]

- 1 Videodisk Equipment
- 2 VTR
- 3 Frame Memory
- 4 Image Compression Equipment
- 5 Computer
- 6 File Server
- 7 Computer